

```

RESULT 10
AX534815
LOCUS AX534815 3221 bp DNA linear PAT 22-NOV-2002
DEFINITION Sequence 53 from Patent WO02068623.
ACCESSION AX534815
VERSION AX534815.1 GI:25261346
KEYWORDS
SOURCE
ORGANISM Aspergillus niger
Aspergillus niger
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE
1
AUTHORS Edens L., dijk Van A.A., Krubasik, P., Albermann, K., Stock, A.,
Kimpel, E., Klugbauer, S., Wagner, C., Fritz, A., Gustedt Von, W.,
Heinrich, O., Maier, D., Spreafico, F., Folkers, U., Hopper, S.,
Kemmer, W., Tan, P., Stiebler, J. and Albang, R.
TITLE Novel genes encoding novel proteolytic enzymes
JOURNAL Patent: WO 02068623-A 53 06-SEP-2002;
DSM N.V. (NL)
FEATURES
Location/Qualifiers
source
1..3221
/organism="Aspergillus niger"
/mol_type="genomic DNA"
/db_xref="taxon:5061"
BASE COUNT 739 a 773 c 869 g 840 t
ORIGIN
Query Match 6.5%; Score 107.8; DB 6; Length 3221;
Best Local Similarity 48.5%; Pred. No. 1e-16;
Matches 449; Conservative 0; Mismatches 437; Indels 39; Gaps 4;
QY 668 TCAGTCTATGACGGAGAGCTATGGAGGCACTATGCTCTGCTGATCTTCAATCTTT 727
Db 1522 TCAGTATATGACGGAGTATATGGGACGGTATGACGGGATACAGGCGCTCTTTC 1581
QY 728 ACAGCAGAAATGAGAGAAATGCCAAGGTAGTGTAAAT---GGTGTACGCTTAATTTCA 784
Db 1582 AGGATGATGAGAGGATGAGGTGGGAGGTAGCACCAGGAGAGATCAATTCG 1641
QY 785 ACTCTCTGGGAATTAATTAACGGCATCATCGACGAGCGATCCAGGCGCTTACTACCTG 844
Db 1642 ATACGCTGGCATATCATAGGTGTGTGATTTACTGTGAGGTCCCTCTGCTCCCTG 1701
QY 845 AATTGCTGTGAACATACCTACGATCAAGGCTGTCAACGAGACCGCTTACACTACA 904
Db 1702 AGCAGCGGTATAACAATACGTATGGATCGAGGATCAATCGACGCTTACGACCGG 1761
QY 905 TGAAGTTTGCCAAACCAATGCCAATGGTGTCCAGGATTTGATTTCCACCTG----- 956
Db 1762 CTATGGATAGTTGAGCAGCGCTGGCGGTGCGAGGATATGATCGAGTGTCCGATG 1821
QY 957 -----CAACAGACAAACCGCACCGCATTAGCTACTACGCTTCTCGG 1881
Db 1822 CTGGCGAGCTCGAGATCCCTCATGTATGGGCAATGAGACGGTAAATAGCATCTCG 1900
QY 1001 CCGAGGCCAACCATGTGCGAGGACATGTTGAGGGGCCATACCTACGCCCTTCTGTC 1060
Db 1882 AGGAGGCTCGGACTACTGTTGCGGGAGATCAAGAGCCTGTATACGATACCTCGGGC 1941
QY 1061 GTGTGTGTATGATATTCGGATCCATATGATGACCGGATCCCGCAAGTTATTACACA 1120
Db 1942 GAGATACTACGATACGCAATTCACGCCGATGACGCTCTCTGCTTACTTCTGCTG 2001
QY 1121 AATTCTGGCAAGGACTCTGTATGAGCGGTATCGGCGTCAACATCACTACACCGAGT 1180
Db 2002 GGTTCTTGAATCGCCCATGGGTGCAAGGCACTTTGGGGTCCCGGTGAATACCATGT 2061
QY 1181 CCAATATGAGCTACTACGCTTTCCAGCAAAACAGCGGACT-----TTGCTGCGC 1231
Db 2062 CGTCAGAGGAGTGGGAAACAGTTTCGCTCGAGGCGGATATCCCGGAATGATCCCC 2121

```

```

QY 1232 CCAACTTCATCGAAGACCTCGAGGATCCCTTGTCTCTCCCGTGGCTGCTCCCTCATCT 1291
Db 2122 CGGGAATGATCGGGATATGGATACTTGTCTTGTGATCCGGTGTCAAGTGGGTATGGTAT 2181
QY 1292 ATGGCGACCGGATATACATCTGCACTGGTTCGGCGGTACAGCCGTTTCCCTCGCTGCGA 1351
Db 2182 ATGGGACCGGACTATGCTTGTCCGTGGCGCGGGGAAAGATCTCAGCTGCTGGTGG 2241
QY 1352 ACTACTCCAAAGCCCGCAGTTCGAAAGCGCAGGATACACGCCCTGAAAGTCAACGGCG 1411
Db 2242 AGTACGAGGATGCGGAGAAGTTCGTGCTGGTATGCCAAAGTGCAGACGAAGTCAAT 2301
QY 1412 TCGATATGGG---AACTCGCAGTATGTAATTTCTCTTCACTCGCTCTATGAGG 1468
Db 2302 CCTACGTTGGGGTCTAGTAAGGCAGTATGGAATCTCTGTTACGGGTGCTTTTCAAG 2361
QY 1469 CAGGCCATGAATCCCATATACACGACCATCGCTCCCTGCAATTTGTTAACGGGACTA 1528
Db 2362 CGGSCCATGAGTGCCCATTTATCAGCCGAAACGGGTATGAGATTTTAAATCGCGTC 2421
QY 1529 TCTTCGTTGGGATATCGCAGAGG 1553
Db 2422 AGTTAATTGGATATTCGACGGG 2446

RESULT 11
LOCUS A75535 2503 bp DNA linear PAT 15-OCT-1999
DEFINITION Sequence 2 from Patent WO9400579.
ACCESSION A75535
VERSION A75535.1 GI:6065609
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 2503)
AUTHORS Fleer, R. and Fournier, A.
TITLE MODIFIED KLUYVEROMYCES YEASTS, THEIR PREPARATION AND USE
JOURNAL Patent: WO 9400579-A 2 06-JAN-1994;
RHONE POULENC RORER SA (FR); FLEER REINHARD (FR)
FEATURES
Location/Qualifiers
source
1..2503
/organism="unidentified"
/mol_type="genomic DNA"
/strain="Kluyveromyces lactis"
/db_xref="taxon:32644"
387..1862
/gene="Kl.PRC1"
387..1862
/gene="Kl.PRC1"
387..1862
/codon_start=1
/product="Gene de la protease C de K.lactis"
/protein_id="CAB58521.1"
/db_xref="GI:6065610"
/translation="MWSIKFLSLYLWLSVTLAISLNAVDSFLPSNFDGNNLIEDHE
RNPNDPVSFSSNDIDAYSILRKLPLKSLGVDTVKQMSGYLDYQSKHFFWFFES
TANPNDPVSILWLNCGPGCSSFVCLFELGPSSIGADLKEINYSWNSASVFLPDQ
PVGFRSYGSKYSTDDAAKDYVFLDLFFERPHLRNDFHLSGESYAGHYLFKIA
HEIAVHAEDSSSNLSVLNGFTDPLTQYQYEPMACGGGYPAVLEPDCDMNR
NLPLCLSLDRCKYKSHSVFSGVLADRYCEQQTITGVYEKGRNPYDIPSKCEARDSDG
CYBEIYISYLNQEEVQRALGTVDSFGQSDVGIGFAFTGDPGSPFPHQYVALLD
QDNVLIYAGDKDYICNWLGNLAWTEKLEWRYNEFYKKVLRVTKWSEETDEITGTS
YGPLTIRIYDAGHMVPHQDPENSLQWNSWNIQAKRSRI."
BASE COUNT 703 a 475 c 517 g 808 t
ORIGIN
Query Match 5.3%; Score 86.4; DB 6; Length 2503;
Best Local Similarity 57.4%; Pred. No. 3.7e-11;
Matches 156; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 199 GTCAATCTCTCTCGATATGTCGACACTCTCCCGAGTCCCATACCTCTCTCTGTTTC 258
Db 624 GTGAAACAATGTCGGGATATTAGATTACAGGACTCAAAACACTCTTTTATTGGTTT 683

```

Mon Nov 24 13:41:03 2003

```

REFERENCE 1 (bases 1 to 2002)
AUTHORS Yaver, D.S. and Thompson, S.A.
TITLE Gene encoding carboxypeptidase of aspergillus niger
JOURNAL Patent: US 5594119-A 3 14-JAN-1997;
FEATURES
    Location/Qualifiers
    source 1..2002
    /organism="unknown"
BASE COUNT 416 a 591 c 505 g 490 t
ORIGIN
Query Match 5.0%; Score 84; DB 6; Length 2002;
Best Local Similarity 55.5%; Pred. No. 1.5e-10; Indels 0; Gaps 0;
Matches 161; Conservative 0; Mismatches 129;
223 GACACCTCTCCCGAGTCCCATACCTTCTCTGTTTCTTCGAAGCCAGACATAACCCAGAA 282
    |||
617 GACAAACGAGAACGACAAAGCATCTGTTCTACTGTTCTTCGAGTCTCGCAATGACCCGAG 676
    |||
283 ACTGCACCTATACATTTGTGTTGAATGGTGGCCCTGGAGCGGATTTTGTATCGGTCTC 342
    |||
677 AATGACCTGTGTTCTGCTGAGCGGTGGCCCTGGATCTCTTCCCTCACCGGTCTT 736
    |||
343 TTCGAAGAGTGGGCCCTTGCATCTCAATTCGACTTTTGATGACTATACCAACCTCAC 402
    |||
737 TTCATGGAGCTCGGCCCTAGCAGCATCAACAAGAGATCCAGCCGGTCTACAACGACTAC 796
    |||
403 TCGTGGACGAGGTCTCCAAATTTACTATCTCTGCCAGCCATTTGGGAGTGGGTTCCTCA 462
    |||
797 GCTTGGAACTCCACGCGTCCGTGATCTCTCTTGACGACCTGTCAACGTTCGGTACTCT 856
    |||
463 TATAGTGTAGCGTTGTGTTGAGTGGTTCATTAACCTGTAACCTGGGTGCTCGA 512
    |||
857 TACAGCACTCTGCTGTCTGAGCGACACCCGTGCTGCTGGCAGGACGCTCA 906
    |||

RESULT 14
174375 174375 2002 bp DNA linear PAT 03-APR-1998
LOCUS Sequence 3 from patent US 5688663.
DEFINITION I74375
ACCESSION I74375
VERSION I74375.1 GI:3010516
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2002)
AUTHORS Yaver, D.Sue. and Thompson, S. Ann.
TITLE Gene encoding carboxypeptidase of Aspergillus niger
JOURNAL Patent: US 5688663-A 3 18-NOV-1997;
FEATURES
    Location/Qualifiers
    source 1..2002
    /organism="unknown"
BASE COUNT 416 a 591 c 505 g 490 t
ORIGIN
Query Match 5.0%; Score 84; DB 6; Length 2002;
Best Local Similarity 55.5%; Pred. No. 1.5e-10; Indels 0; Gaps 0;
Matches 161; Conservative 0; Mismatches 129;
223 GACACCTCTCCCGAGTCCCATACCTTCTCTGTTTCTTCGAAGCCAGACATAACCCAGAA 282
    |||
617 GACAAACGAGAACGACAAAGCATCTGTTCTACTGTTCTTCGAGTCTCGCAATGACCCGAG 676
    |||
283 ACTGCACCTATACATTTGTGTTGAATGGTGGCCCTGGAGCGGATTTTGTATCGGTCTC 342
    |||
677 AATGACCTGTGTTCTGCTGAGCGGTGGCCCTGGATCTCTTCCCTCACCGGTCTT 736
    |||
343 TTCGAAGAGTGGGCCCTTGCATCTCAATTCGACTTTTGATGACTATACCAACCTCAC 402
    |||
737 TTCATGGAGCTCGGCCCTAGCAGCATCAACAAGAGATCCAGCCGGTCTACAACGACTAC 796
    |||
403 TCGTGGACGAGGTCTCCAAATTTACTATCTCTGCCAGCCATTTGGGAGTGGGTTCCTCA 462
    |||

REFERENCE 13
I33983 I33983 2002 bp DNA linear PAT 06-FEB-1997
LOCUS Sequence 3 from patent US 5594119.
DEFINITION I33983
ACCESSION I33983
VERSION I33983.1 GI:1824774
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
259 TTCGAAGCCAGACATACCAGAACTGCACCTATACATTTGTTGTTGAATGGTGGCCCT 318
    |||
684 TTTGAGTCTAGAAATGACCCAGAGATGACCCAGTATGATGTTTAAACGGTGTGCTT 743
    |||
319 GGAAGCGATTCTTTGATCGGTCTCTTCGAAGAGTGGGCCCTTGGCCATGTCATTCGACT 378
    |||
744 GGCTGTTCTCTTTCGTCGGTCTTTCTTTGAATGGGACCTTCTTCTATAGAGCTGAT 803
    |||
379 TTTGATGACTACATCAACCCCTCACTCGTGGAGCGAGGTCTCCAAATTTACTATTCCTGTC 438
    |||
804 TTGAACCCATTATACCCCTACTCTTGGAAATCCAAACGCTTCTGTGATATTCCTAGAT 863
    |||
439 CAGCCATTGGGAGTCGGCTTTTCATATAGTCA 470
    |||
864 CAGCCTGTGGTGTGGGTCTCTCATACGGTGA 895
    |||

RESULT 12
I70282 I70282 2503 bp DNA linear PAT 03-APR-1998
LOCUS Sequence 3 from patent US 5679544.
DEFINITION I70282
ACCESSION I70282
VERSION I70282.1 GI:3006417
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2503)
AUTHORS Fleer, R., Fournier, A. and Yeh, P.
TITLE Modified Kluyveromyces yeasts, their preparation and use
JOURNAL Patent: US 5679544-A 3 21-OCT-1997;
FEATURES
    Location/Qualifiers
    source 1..2503
    /organism="unknown"
BASE COUNT 703 a 475 c 517 g 808 t
ORIGIN
Query Match 5.2%; Score 86.4; DB 6; Length 2503;
Best Local Similarity 57.4%; Pred. No. 3.7e-11; Indels 0; Gaps 0;
Matches 156; Conservative 0; Mismatches 116;
199 GTCAAATCTACTCTCGATATGTCGACACCTCTCCGAGTCCCATACCTTCTTCTGTTTC 258
    |||
624 GTGAAACATGTCGGGATATTTAGATTACGAGACTCAAAACACTTCTTTTATTGTTT 693
    |||
259 TTCGAAGCCAGACATACCAGAACTGCACCTATACATTTGTTGTTGAATGGTGGCCCT 318
    |||
684 TTTGAGTCTAGAAATGACCCAGAGATGACCCAGTATGATGTTTAAACGGTGTGCTT 743
    |||
319 GGAAGCGATTCTTTGATCGGTCTCTTCGAAGAGTGGGCCCTTGGCCATGTCATTCGACT 378
    |||
744 GGCTGTTCTCTTTCGTCGGTCTTTCTTTGAATGGGACCTTCTTCTATAGAGCTGAT 803
    |||
379 TTTGATGACTACATCAACCCCTCACTCGTGGAGCGAGGTCTCCAAATTTACTATTCCTGTC 438
    |||
804 TTGAACCCATTATACCCCTACTCTTGGAAATCCAAACGCTTCTGTGATATTCCTAGAT 863
    |||
439 CAGCCATTGGGAGTCGGCTTTTCATATAGTCA 470
    |||
864 CAGCCTGTGGTGTGGGTCTCTCATACGGTGA 895
    |||

```

Db 797 GCTTGAACCTCAACGGTCCGTGATCTTCCTTGACAGCCTGTCAACGTCGGTTACTCT 856
QY 463 TATAGTGATACGGTTGATGGTCCATTAAACCTGTAACTGGGTGCTGCA 512
Db 857 TACAGCAACTCTGCTGTACAGCACCGTGTGCTGCTGGCAAGACGTCTA 906

RESULT 15
I77239
LOCUS I77239 2002 bp DNA linear PAT 03-APR-1998
DEFINITION Sequence 3 from patent US 5693510.
ACCESSION I77239
VERSION I77239.1 GI:3013393
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2002)
AUTHORS Xaver, D. Sue. and Thompson, S. Ann.
TITLE Gene encoding carboxypeptidase of Aspergillus niger
JOURNAL Patent: US 5693510-A 3 02-DEC-1997;
FEATURES
source Location/Qualifiers
1..2002
/organism="unknown"

BASE COUNT 416 a 591 c 505 g 490 t
ORIGIN

Query Match 5.0%; Score 84; DB 6; Length 2002;
Best Local Similarity 55.5%; Pred. No. 1.5e-10;
Matches 161; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 223 GACACCTCTCCGAGTCCCATACCTTCTTGTTCTTCGAAGCCAGACATAACCCAGAA 282
Db 617 GACAAAGAGACGACGACGATCTGTTCTTCTGAGTCTCGCAATGACCCCGAG 676
QY 283 ACTGCACCTATCACATTGTTGTTGAATGGTGGCCCTGGAAGCGATTCTTTGATCGGTCTC 342
Db 677 AATGACCTTGTTCTGTGGTGAACGGTGGCCCTGGATGCTTTCCTCACCCTGCTT 736
QY 343 TTCGAAGAGTTGGGCCCTTGGCCATGTCATTCGACTTTTGTGACTACATCAACCTTCAC 402
Db 737 TTCATGGAGCTCGGCCCTAGCAGCATCAACAAGAAGATCCAGCCGGTCTACAAGACTAC 796
QY 403 TCGTGAACGAGGTCCTCAATTACTATTCCTGTCGCCAGCATTCGGAGTCGGCTTTTCA 462
Db 797 GCTTGAACCTCAACGCTCGGTGATCTTCCTTGACAGCCTGTCACGTCGGTTACTCT 856
QY 463 TATAGTGATACGGTTGATGGTCCATTAAACCTGTAACTGGGTGCTGCA 512
Db 857 TACAGCAACTCTGCTGTACAGCACACCGTGTGCTGCTGGCAAGGACGTCTA 906

Search completed: November 21, 2003, 13:23:24
Job time : 6074 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:45:55 ; Search time 74.5 Seconds
(without alignments)
11555.215 Million cell updates/sec

Title: US-09-712-338-1
Perfect score: 555
Sequence: 1 atgcgtggctacgaattctt.....ccagtgtgtgatgcatag 1668

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues
Word size: 1

Total number of hits satisfying chosen parameters: 1660916

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame-n2p.model -DEV=xl
-Q/cn2_1/USPTO.spool_p/US09712338/runat 17112003 170226 10771/app query.fasta 1.1863
-DB=SPTRMBL 23 -QFMT=fastan -SUFFIX=oli.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @CGN 1 1 127 @runat 17112003 170226 10771 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SPTRMBL 23:*
1: sp archaea:*
2: sp bacteria:*
3: sp fungi:*
4: sp human:*
5: sp invertebrate:*
6: sp mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	Match	Length	DB ID	Description
1	454	81.8	555	3 Q96VZ9	Q96VZ9 aspergillus

SUMMARIES

2	10	1.8	460	3	Q12569
3	9	1.6	67	2	Q937G9
4	9	1.6	145	16	Q8RDZ1
5	9	1.6	147	2	Q936V3
6	9	1.6	164	5	Q9GRR4
7	9	1.6	290	16	Q8D686
8	9	1.6	375	17	Q8TU72
9	9	1.6	405	2	O66024
10	9	1.6	405	2	P71195
11	9	1.6	405	2	Q91UM7
12	9	1.6	405	2	O56451
13	9	1.6	405	2	Q9F3W1
14	9	1.6	468	10	Q9SHU3
15	9	1.6	469	10	Q9FH06
16	9	1.6	473	10	Q9FH05
17	9	1.6	479	10	Q9WAR8
18	9	1.6	495	2	Q48384
19	9	1.6	1535	5	Q9VFF0
20	9	1.6	2105	5	Q17679
21	8	1.4	51	8	Q957T5
22	8	1.4	63	8	Q957T1
23	8	1.4	63	8	Q957S9
24	8	1.4	76	4	Q96GB0
25	8	1.4	91	16	Q9R48
26	8	1.4	91	16	Q8NUN8
27	8	1.4	99	17	Q971Q8
28	8	1.4	114	17	Q8TJ61
29	8	1.4	116	15	Q75691
30	8	1.4	166	17	O59236
31	8	1.5	169	10	Q9ASH5
32	8	1.5	172	16	Q8XX16
33	8	1.5	184	17	O27824
34	8	1.5	191	10	Q8LJM1
35	8	1.5	203	16	Q8DEH0
36	8	1.4	210	17	Q9HMH2
37	8	1.5	249	4	Q8H8T6
38	8	1.4	259	2	Q9S3T5
39	8	1.4	263	2	Q9ZFF5
40	8	1.4	264	10	Q9LXY6
41	8	1.4	264	10	Q8GVD2
42	8	1.4	270	16	Q9RRC8
43	8	1.4	281	16	Q9KS32
44	8	1.5	290	16	Q8GZ29
45	8	1.5	326	2	O68878

ALIGNMENTS

RESULT 1
Q96VZ9
ID Q96VZ9 PRELIMINARY; PRT; 555 AA.
AC Q96VZ9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Carboxypeptidase S1 precursor (EC 3.4.16.6).
GN CPl.
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OC NCBI_TaxID=5062;
FN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TK3;
RA van den Broek P.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF394242; AAK77166.1; -
DR InterPro: IPR001563; Serine carboxypeptidase.
DR Pfam: PF00450; serine carboxypeptidase; 1.
DR PRINTS: PR00724; CRBOXPTASEC.
DR ProDom: PD001189; Serine carboxypeptidase; 1.
DR PROSITE: PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase; Hydrolase; Signal.

QY	901	TACATGAAGTTTCCCAACCAATGCCAAATGGTTGCCAGATTAGCTACGCTTCTGGCCGCGAAGCCACCAACATGTGC	1020
DB	301	TyrMetLysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSerThrCysLys	320
QY	961	CAGACAAACCGCACCGCATTTAGCTACGCTTCTGGCCGCGAAGCCACCAACATGTGC	1020
DB	321	GlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCys	340
QY	1021	AGGACAAATGTTGAGGGCCATATACGCTTGTGCTGT	

DR EMBL; D16519; BAA03966.1; -;
 DR HSSP; P00729; IYSC.
 DR MEROPS; S10.010; -;
 DR InterPro; IPR001563; Serine carboxypept.
 DR InterPro; IPR000379; Ser esters site.
 DR Pfam; PF00450; serine carboxypept; 1.
 DR PRINTS; PR00724; CRBOXYPIASEC.
 DR ProDom; PD001189; Serine carboxypept; 1.
 KW Carboxypeptidase; Hydrolase.
 FT CHAIN 52 460 MATURE CARBOXYPEPTIDASE Z.
 SQ SEQUENCE 460 AA; 50825 MW; CE3A410764548B6 CRC64;

Alignment Scores:
 Pred. No.: 0.654 Length: 460
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.80% Indels: 0
 DB: 3 Gaps: 0

US-09-712-338-1 (1-1668) x Q12569 (1-460)

QY 1456 CCGCTCTATGAGCGAGCCATGAAGTCCCA 1485
 |||||
 Db 430 ArgValTyGluAlaGlyHisGluValPro 439

RESULT 3
 Q937G9
 ID Q937G9 PRELIMINARY; PRT; 67 AA.
 AC Q937G9;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Putative transposition protein (Fragment).
 GN TniQ.
 OS Klebsiella sp. L513-39.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Klebsiella.
 OX NCBI_TaxID=143776;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LS13-39; TRANSPOSON=Tn5056;
 RX MEDLINE=21604134; PubMed=11763242;
 RA Mindlin S.Z., Kholodil G.Y., Gorlenko Z.M., Minakhina S.V.,
 RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
 RA Yurieva O.V., Nikiforov V.G.;
 RT "Mercury resistance transposons of Gram-negative environmental
 bacteria and their classification."
 RL Res. Microbiol. 152:811-822(2001).
 DR EMBL; AJ302770; CAC82957.1; -;
 FT NON-TER 1
 SQ SEQUENCE 67 AA; 7547 MW; 3EECC0A1149C62D9 CRC64;

Alignment Scores:
 Pred. No.: 9.83 Length: 67
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x Q937G9 (1-67)

QY 262 GAAGCCAGACATAACCCAGAACTGCA 288
 |||||
 Db 3 GluAlaArgHisAsnProGluThrAla 11

RESULT 4
 Q8RDZ1
 ID Q8RDZ1 PRELIMINARY; PRT; 145 AA.
 AC Q8RDZ1;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Integral membrane protein.
 GN FNI350.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
 OC Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RX MEDLINE=21886394; PubMed=11889109;
 RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fonstein M., Kyrpides N., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 nucleatum strain ATCC 25586."
 RL J. Bacteriol. 184:2005-2018(2002).
 DR EMBL; AE01639; AAL95546.1; -;
 KW Complete proteome.
 SQ SEQUENCE 145 AA; 15902 MW; 7DB20993CC1C6F0E CRC64;

Alignment Scores:
 Pred. No.: 8.82 Length: 145
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 16 Gaps: 0

US-09-712-338-1 (1-1668) x Q8RDZ1 (1-145)

QY 761 TTAATGGTTCAGCTTAATTTCAACT 787
 |||||
 Db 131 LeuMetValPheSerLeuIleSerThr 139

RESULT 5
 Q936V3
 ID Q936V3 PRELIMINARY; PRT; 147 AA.
 AC Q936V3;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Putative transposition protein TniQ (Fragment).
 GN TniQ.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BW20; TRANSPOSON=Tn5053v4;
 RX MEDLINE=21604134; PubMed=11763242;
 RA Mindlin S.Z., Kholodil G.Y., Gorlenko Z.M., Minakhina S.V.,
 RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
 RA Yurieva O.V., Nikiforov V.G.;
 RT "Mercury resistance transposons of Gram-negative environmental
 bacteria and their classification."
 RL Res. Microbiol. 152:811-822(2001).
 DR EMBL; AJ296313; CAC82974.1; -;
 FT NON-TER 1
 SQ SEQUENCE 147 AA; 16500 MW; 24636FA7A1489ABD CRC64;

Alignment Scores:
 Pred. No.: 8.81 Length: 147
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x Q936V3 (1-147)

Alignment Scores: 8 Length: 290
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.62% Gaps: 0
DB: 16

US-09-712-338-1 (1-1668) x Q8D686 (1-290)

QY 227 CCTCTCCGAGTCCCATACCTCTCTCT 253
|||||
Db 161 ProLeuProSerProLeuProSerSer 169
|||||

RESULT 8
Q8TU72 PRELIMINARY; PRT; 375 AA.
ID Q8TU72
AC Q8TU72
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE 3-isopropylmalate dehydrogenase.
GN LEUB OR MA0201.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearrellan K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talanas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.I., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010678; AA003654.1; -;
DR InterPro; IPR001804; Isohd.
DR Pfam; PF00180; isodh; 1.
DR PROSITE; PS00470; IDH_IMDH; 1.
KW Complete proteome.
SQ SEQUENCE 375 AA; 41318 MW; 35646A9B4A005AFC CRC64;

Alignment Scores: 7.72 Length: 375
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.65% Gaps: 0
DB: 17

US-09-712-338-1 (1-1668) x Q8TU72 (1-375)

QY 1272 GGGGAGAGGATCTCTCGAGTCC 1246
|||||
Db 147 GlycylserineAspLeuLeuGluVal 155
|||||

RESULT 9
Q66024 PRELIMINARY; PRT; 405 AA.
ID Q66024
AC Q66024
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)

QY 262 GRAGCCAGACATACCCAGAACTGCA 288
|||||
Db 83 GluAlaArgHisAsnProGluThrAla 91
|||||

RESULT 6
Q9GRR4 PRELIMINARY; PRT; 164 AA.
ID Q9GRR4
AC Q9GRR4
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Conserved hypothetical protein L5882.03.
GN L5882.03.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Masuy D., Purnelle B., Goffeau A., Ivens A.C., Quail M.,
RA Rajandream M.A., Barrell B.G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
GN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; AL445943; CAC14321.1; -;
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 17935 MW; F60D849331340AE7 CRC64;

Alignment Scores: 8.67 Length: 164
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.65% Gaps: 0
DB: 5

US-09-712-338-1 (1-1668) x Q9GRR4 (1-164)

QY 185 TCGCAGCGCTCTGCCCCGGTTC 159
|||||
Db 154 SerGlnThrProSerAlaProGlySer 162
|||||

RESULT 7
Q8D686 PRELIMINARY; PRT; 290 AA.
ID Q8D686
AC Q8D686
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN V20650.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB016810; AAC07593.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 290 AA; 33267 MW; AC8FF330BE7F0562 CRC64;

```

DE TniQ.
GN TniQ.
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OG Plasmid pPB.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OX;
RX MEDLINE=98148002; PubMed=9479042;
RA Reniero D., Mozzon E., Galli E., Barbieri P.;
RT "Two aberrant mercury resistance transposons in the Pseudomonas
stutzeri plasmid pPB.";
RL Gene 208:37-42(1998).
DR EMBL; U81032; AAC38239.1; -.
KW Plasmid.
SQ SEQUENCE 405 AA; 45761 MW; EBB5C33CCE6C7CFF CRC64;

Alignment Scores:
Pred. No.: 7.64 Length: 405
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x O66024 (1-405)
QY 262 GAAGCCAGACATAACCCAGAACTGCA 288
Db 341 GluAlaArgHisAsnProGluThrAla 349

RESULT 10
ID P71195 PRELIMINARY; PRT; 405 AA.
AC P71195;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Plasmid R751, complete sequence.
OS Enterobacter aerogenes (Aerobacter aerogenes).
OG Plasmid R751.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Enterobacter.
OX NCBI_TaxID=548;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98428622; PubMed=9753548;
RA Thorsted P.B., Macartney D.P., Akhtar P., Haines A.S., Ali N.,
Davidson P., Stafford T., Pocklington M.J., Pansegrau W.,
Wilkins B.M., Lanka E., Thomas C.M.;
RT "Complete sequence of the IncPbeta plasmid R751: implications for
evolution and organisation of the IncP backbones.";
RN J. Mol. Biol. 282:969-990(1998).
RP SEQUENCE FROM N.A.
RA Thomas C.M.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Haines A.S., Thomas C.M.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U67194; AAC64465.1; -.
KW Plasmid.
SQ SEQUENCE 405 AA; 45567 MW; 0F0749CDA19244F7 CRC64;

Alignment Scores:
Pred. No.: 7.64 Length: 405
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0

```

```

DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x P71195 (1-405)
QY 262 GAAGCCAGACATAACCCAGAACTGCA 288
Db 341 GluAlaArgHisAsnProGluThrAla 349

RESULT 11
Q91UM7 PRELIMINARY; PRT; 405 AA.
ID Q91UM7;
AC Q91UM7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE TniQ protein.
GN TniQ.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSB102
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=mercury resistance transposon Tns718;
RA Schneider S., Keller M., Droegge M., Lanka E., Puehler A.,
Selbitschka W.;
RT "The genetic organization and evolution of the broad-host-range
mercury resistance plasmid pSB102 isolated from a microbial population
residing in the rhizosphere of alfalfa.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ304453; CAC79210.1; -.
KW Plasmid.
SQ SEQUENCE 405 AA; 45711 MW; 2396F2C89A2DE7B2 CRC64;

Alignment Scores:
Pred. No.: 7.64 Length: 405
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x Q91UM7 (1-405)
QY 262 GAAGCCAGACATAACCCAGAACTGCA 288
Db 341 GluAlaArgHisAsnProGluThrAla 349

RESULT 12
Q56451 PRELIMINARY; PRT; 405 AA.
ID Q56451;
AC Q56451;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE Mercury resistant transposon TNS053 (Central Asia).
GN TniQ.
OS Xanthomonas sp.
OG Plasmid RPI.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=29446;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93253772; PubMed=8387603;
RA Kholodii G.Y.A., Yurieva O.V., Lomovskaya O.L., Gorlenko Zh.M.,
Mindlin S.Z., Nikiforov V.G.;
RT "Tns053, a mercury resistance transposon with integron's ends.";
RN J. Mol. Biol. 230:1103-1107(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96130850; PubMed=8594337;

```


RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones.";

RL DNA Res.: 7:31-63 (2000).
 DR EMEL; AB023032; BABL0196.1; -.
 DR HSP; P08819; 1WHT.
 DR MEROPS; S10.005; -.
 DR InterPro; IPR001563; Serine_carbpept.
 DR Pfam; PF00450; serine_carbpept; 1.
 DR PRINTS; PR00724; CREOXYPTASEC.
 DR ProDom; PD001189; Serine_carbpept; 1.
 DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
 KW Carboxypeptidase.
 SQ SEQUENCE 469 AA; 52541 MW; 90357E27B9B64527 CRC64;

Alignment Scores:

Pred. No.:	7.48	Length:	469
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.62%	Indels:	0
DB:	10	Gaps:	0

US-09-712-338-1 (1-1668) x Q9FH06 (1-469)

QY	295	ACATTGTGTTGAATGGTGGCCCTGGA	321
Db	74	ThrLeuTripleuAsnGlyProGly	82

Search completed: November 21, 2003, 17:54:35
 Job time : 83.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:44:24 ; Search time 19 Seconds
(without alignments)
8256.899 Million cell updates/sec

Title: US-09-712-338-1
Perfect score: 555
Sequence: 1 atgcgtggtaacgaattctt.....ccagtgttgatggcatag 1668

Scoring table:

OLIGO	Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0	
Egapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 127863 seqs, 47026705 residues

Word size: 1

Total number of hits satisfying chosen parameters: 255550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US09712338/runat.17112003.170225.1076/app_query.fasta_1.1863
-DB=SwissProt 41 -OPMT=fastan -SUFFIX=oli.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAIRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @CGN 1.1.25 @runat.17112003.170225.10761 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7
```

Database : SwissProt 41.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	1.8	423	1	CPS1_PENJA
2	9	1.6	508	1	YBY9_YEAST
3	8	1.4	36	1	P3BY_ODOSI
4	8	1.4	292	1	YCL5_PYRO
5	8	1.4	356	1	KRV_MOUSE
6	8	1.4	423	1	CBP2_MOUSE
7	8	1.4	436	1	CBP2_HORVU
8	8	1.4	470	1	YSS2_CAEEL
9	8	1.4	474	1	PRTP_MOUSE
10	8	1.4	476	1	CBP2_HORVU
11	8	1.4	480	1	PRTP_HUMAN
12	8	1.4	499	1	CBP1_HORVU
13	8	1.4	510	1	CBP1_ORYGA
14	8	1.4	516	1	CBP2_HORVU
15	8	1.4	523	1	CBP1_PICPA
16	8	1.5	526	1	VSM4_TRYEB
17	8	1.4	532	1	CBPY_YEAST
18	8	1.4	542	1	CBPY_CANAL

19	8	1.4	574	1	YPP3_CAEEL
20	8	1.4	605	1	WSC4_YEAST
21	8	1.5	760	1	EZ_DROME
22	8	1.4	1002	1	CBFY_SCHPO
23	8	1.4	2145	1	CYAA_PODAN
24	8	1.4	4349	1	FAT2_HUMAN
25	7	1.3	25	1	YCX8_ODOSI
26	7	1.3	66	1	PSAE_FREDI
27	7	1.3	70	1	PSAE_CYPAP
28	7	1.3	88	1	YPB4_ECOLI
29	7	1.3	91	1	YLS9_CAEEL
30	7	1.3	110	1	RLA4_YEAST
31	7	1.3	112	1	H34_STYLE
32	7	1.3	113	1	H38_STYLE
33	7	1.3	114	1	H31_STYLE
34	7	1.3	114	1	H33_STYLE
35	7	1.3	114	1	H35_STYLE
36	7	1.3	114	1	H36_STYLE
37	7	1.3	114	1	H37_STYLE
38	7	1.3	114	1	H37_STYLE
39	7	1.3	114	1	ARR2_ECOLI
40	7	1.3	117	1	SMD1_SCHPO
41	7	1.3	124	1	PA25_AGKHP
42	7	1.3	126	1	YE66_AQUAE
43	7	1.3	132	1	RS8_MYCLE
44	7	1.3	133	1	YGF1_ECOLI
45	7	1.3	134	1	H3_CHLRE

ALIGNMENTS

RESULT 1	CPS1_PENJA	STANDARD;	PRT;	423 AA.
ID	CPS1_PENJA			
AC	P34946;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Carboxypeptidase S1 (EC 3.4.16.6)			
OS	Penicillium janthinellum (Penicillium vitale)			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.			
OX	NCBI_TaxID=5079;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=94039747; PubMed=8224168;			
RA	Svendsen I., Hofmann T., Endrizzi J., Remington S.J., Breddam K.;			
RT	"The primary structure of carboxypeptidase S1 from Penicillium janthinellum."			
RL	FEBS Lett. 333:39-43 (1993).			
CC	-!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine or lysine residue.			
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.			
DR	PIR; S38953; S38953.			
DR	HSSP; P00729; 1YSC.			
DR	MEROPS; S10.008; --			
DR	InterPro; IPR000379; Ser esters site.			
DR	InterPro; IPR001563; Serine carboxypeptidase.			
DR	Pfam; PF00450; serine carboxypeptidase.			
DR	PRINTS; PR00724; CRBOXYPTASEC.			
DR	ProDom; PD001189; Serine carboxypeptidase.			
DR	PROSITE; PS00131; CARBOXYPEPT SER SER; 1.			
DR	PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.			
DR	Hydrolase; Carboxypeptidase; Glycoprotein.			
FT	ACT SITE 143 143			POTENTIAL
FT	ACT SITE 340 340			BY SIMILARITY.
FT	ACT SITE 397 397			BY SIMILARITY.
FT	BINDING 343 343			SUBSTRATE.
FT	BINDING 398 398			SUBSTRATE.
FT	DISULFID 8 68			
FT	DISULFID 55 300			
FT	DISULFID 223 246			
FT	DISULFID 230 239			

P25716	caenorhabditis
P38739	saccharomyc
P42124	drosophila
Q13849	schizosacch
Q15133	podospira a
Q9nyq8	homo sapien
P23809	odontella s
P48114	cyamophora
P03853	escherichia
P34334	caenorhabditis
P02400	saccharomyc
P81198	stylyonchia
P81202	stylyonchia
P81195	stylyonchia
P81197	stylyonchia
P81199	stylyonchia
P81200	stylyonchia
P81201	stylyonchia
P81196	stylyonchia
P52144	escherichia
O42661	schizosacch
O42189	agkistrodon
O67446	aquifex aeo
O32997	mycobacteri
P39291	escherichia
P50564	chlamydomon

DR PROSITE; PS00560; CARBOXYPEPT SER_HIS; 1.
KW Hypothetical protein; Hydrolase; Carboxypeptidase.
FT ACT_SITE 219 219 BY SIMILARITY.
FT ACT_SITE 415 415 BY SIMILARITY.
FT ACT_SITE 474 474 BY SIMILARITY.
SQ SEQUENCE 508 AA; 57639 MW; AAB2806C8EE2EBE1 CRC64;
Alignment Scores:
Pred. No.: 2.85 Length: 508
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 1 Gaps: 0
US-09-712-338-1 (1-1668) x YBY9_YEAST (1-508)
QY 439 CAGCCATTGGAGTCGGCTTTTCATAT 465
Db 171 GlnProLeuglyValGlyPheSerTyr 179
RESULT 3
PSBY_ODOSI STANDARD; PRT; 36 AA.
AC P49543;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem II protein Y.
GN (PSBY-A OR PSBY-1 OR YCF32.1) AND (PSBY-B OR PSBY-2 OR YCF32.2).
OS Odontella sinensis (Marine centric diatom).
OG Chloroplast.
OC Eukaryota; Stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Bidulphiophycidae; Eupodiscales; Odontella.
OX NCBI_TaxID=2839;
RN [1]
RP SEQUENCE FROM N.A.
RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pandic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RL Odontella sinensis";
RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC -!- FUNCTION: MANGANESE-BINDING POLYPEPTIDE WITH L-ARGININE
CC METABOLIZING ENZYME ACTIVITY. COMPONENT OF THE CORE OF PHOTOSYSTEM
CC II (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -!- SIMILARITY: BELONGS TO THE PSBY FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z67753; CAA91612.1; -
CC EMBL; Z67753; CAA91666.1; -
CC PIR; S78239; S78239.
KW Photosystem II; Chloroplast; Transmembrane; Thylakoid.
FT DOMAIN 1 6 LUMINAL (POTENTIAL).
FT TRANSMEM 7 23 POTENTIAL.
FT DOMAIN 24 36 STROMAL (POTENTIAL).
SQ SEQUENCE 36 AA; 3921 MW; 22989D3E7DAE2B63 CRC64;
Alignment Scores:
Pred. No.: 45 Length: 36
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 1 Gaps: 0
US-09-712-338-1 (1-1668) x PSBY_ODOSI (1-36)
FT CARBOHYD 200 200 N-LINKED (GLCNAC...);
SQ SEQUENCE 423 AA; 46496 MW; 55F3333608B12729 CRC64;
Alignment Scores:
Pred. No.: 0.296 Length: 423
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 1 Gaps: 0
US-09-712-338-1 (1-1668) x CPS1_PENJA (1-423)
QY 679 ACGGAGGATGAGGAGGACATGTCCT 708
Db 141 ThrGluSerTyrGlyGlyHisTyrGlyPro 150
RESULT 2
YBY9_YEAST STANDARD; PRT; 508 AA.
AC P38109;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative serine carboxypeptidase in ESRI-IRAI intergenic region
DE (EC 3.4.16.-).
GN YBR139W OR YBR1015.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=9437817; PubMed=8091856;
RA Becam A.-M., Cullin C., Gryzbowska E., Lacroute F., Nasr F.,
RA Ozier-Kalogeropoulos O., Palucha A., Slonimski P.P., Zagulski M.,
RA Herbert C.J.;
RT "The sequence of 29.7 kb from the right arm of chromosome II reveals
RT 13 complete open reading frames, of which ten correspond to new
RT genes";
RL Yeast 10:S1-S11(1994).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=95042830; PubMed=7954890;
RA Nasr F., Becam A.-M., Gryzbowska E., Zagulski M., Slonimski P.P.,
RA Herbert C.J.;
RT "An analysis of the sequence of part of the right arm of chromosome
RT II of S. cerevisiae reveals new genes encoding an amino-acid permease
RT and a carboxypeptidase";
RL Curr. Genet. 26:1-7(1994).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X75891; CAA53497.1; -
CC EMBL; Z36008; CAA85097.1; -
CC PIR; S46008; S46008.
CC HSSP; P00729; 1CFY.
CC MROPS; S10.UFW.-
CC SGD; S0000343; YBR139W.
CC InterPro; IPR000379; Ser_estrs_site.
CC InterPro; IPR001563; Serine_carbpept.
CC Pfam; PF00450; serine_carbpept. 1.
CC PRINTS; PR00724; CRBOXYPTASEC.
CC ProDom; PD001189; Serine_carbpept. 1.
CC PROSITE; PS00131; CARBOXYPEPT_SER_1.
CC PROSITE; PS00131; CARBOXYPEPT_SER_1.

Qy 34 TTGGTTCAGCCAGTTGGCCCTT 57
 Db 13 LeuValAlaAlaSerTrpAlaLeu 20

RESULT 4

YC15_PVRHO ID YC15_PVRHO STANDARD; PRT; 292 AA.
 AC OS8968;
 DT 13-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical ABC transporter permease protein PH1215.
 GN PH1215 OR PHK039.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawanabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
 RA Roki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).
 CC -!- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT
 CC OF THE SUBSTRATE ACROSS THE MEMBRANE.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: BELONGS TO THE BINDING-PROTEIN-DEPENDENT TRANSPORT
 CC SYSTEM PERMEASE FAMILY. MALFG SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AP000005; BAA30315.1; -;
 DR F1R; A71065; A71065.
 DR InterPro; IPR000515; BPD transp.
 DR Pfam; PF00528; BPD transp; 1.
 DR PROSITE; PS00402; BPD_TRANSP_INN_MEMBR; 1.
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 10 30 POTENTIAL.
 FT TRANSMEM 73 93 POTENTIAL.
 FT TRANSMEM 106 126 POTENTIAL.
 FT TRANSMEM 160 180 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 261 281 POTENTIAL.
 SQ SEQUENCE 292 AA; 33407 MW; 5B968812C0B89F68 CRC64;

Alignment Scores:
 Pred. No.: 31.1 Length: 292
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.44% Indels: 0
 DB: 1 Gaps: 0

US-09-712-338-1 (1-1668) x YC15_PVRHO (1-292)

Qy 755 GTAGTGTAAATGGTTCAGCTTA 778
 Db 219 ValValLeuMetValPheSerLeu 226

RESULT 5

CKRV_MOUSE ID CKRV_MOUSE STANDARD; PRT; 356 AA.
 AC P51676;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE C-C chemokine receptor 1-like protein 1 (Macrophage inflammatory
 DE protein-1 alpha receptor-like 1).
 GN CCR1L1 OR CMKBR1L1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=95340546; PubMed=7542241;
 RA Gao J.-L., Murphy P.M.;
 RT "Cloning and differential tissue-specific expression of three mouse
 RT beta chemokine receptor-like genes, including the gene for a
 RT functional macrophage inflammatory protein-1 alpha receptor.";
 RL J. Biol. Chem. 270:17494-17501(1995).
 CC -!- FUNCTION: PROBABLE RECEPTOR FOR A C-C TYPE CHEMOKINE.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: DETECTED IN THE SPLEEN, LIVER AND LEUKOCYTES.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U28405; AAA89154.1; -;
 DR F1R; I49340; I49340.
 DR MGD; MGI:104617; Ccr1l1.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PROSITE; PS00237; G-PROTEIN RECP F1_1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECP F1_2; 1.
 KW G-protein coupled receptor; Transmembrane.
 FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 33 60 1 (POTENTIAL).
 FT DOMAIN 61 67 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 68 92 2 (POTENTIAL).
 FT DOMAIN 93 108 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 109 130 3 (POTENTIAL).
 FT DOMAIN 131 147 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 148 172 4 (POTENTIAL).
 FT DOMAIN 173 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 224 5 (POTENTIAL).
 FT DOMAIN 225 240 6 (POTENTIAL).
 FT TRANSMEM 241 265 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 266 282 7 (POTENTIAL).
 FT TRANSMEM 283 306 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 307 356 EXTRACELLULAR (POTENTIAL).
 FT DISULFID 107 184 BY SIMILARITY.
 SQ SEQUENCE 356 AA; 40934 MW; 58C01ABA8D7D4B06 CRC64;

Alignment Scores:
 Pred. No.: 30.1 Length: 356
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.44% Indels: 0
 DB: 1 Gaps: 0

US-09-712-338-1 (1-1668) x CKRV_MOUSE (1-356)

Alignment Scores:		
Pred. No.:	29.2	423
Score:	8.00	8
Percent Similarity:	100.00%	0
Conservative:		

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 1 Gaps: 0

US-09-712-338-1 (1-1668) x CBP2_WHEAT (1-423)
QY 298 TTGTGTTGAATGTCGCCCTGGA 321
  |||||
Db 57 LeuTriPLeuAsnGlyGlyProGly 64

RESULT 7
CP22_HORVU
ID_CP22_HORVU STANDARD; PRT; 436 AA.
AC P55748;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase II-2 precursor (EC 3.4.16.6) (CP-MII.2)
DE (Fragment).
GN CXP2-2.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=cv. Alexis; TISSUE=Grain;
RX MEDLINE=94336715; PubMed=7520177;
RA Dal Degun F., Rocher A., Cameron-Mills V., von Wettstein D.;
RT "The expression of serine carboxypeptidases during maturation and
  germination of the barley grain.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:8209-8213 (1994).
CC -!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine
  or lysine residue.
CC -!- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS
  COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE GERMINATING EMBRYO. LOW
  LEVELS IN THE DEVELOPING ALEURONE AND EMBRYO. ALSO FOUND IN THE
  ROOTS AND SHOOTS OF THE GROWING SEEDLING.
CC -!- PTM: THE LINKER PEPTIDE IS ENDOPEPTIDOLYTIALLY EXCISED DURING
  ENZYME MATURATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; X78878; CAB59202.1; -.
DR HSSP; P08819; 1WHT.
DR InterPro; IPR000379; Ser esters site.
DR InterPro; IPR001563; Serine carboxpept.
DR Pfam; PF00450; serine carboxpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydrolase; Carboxypeptidase; Zymogen; Multigene family.
FT NON TER 1
FT CHAIN <1 256
FT PROPEP 257 270
FT CHAIN 271 436
FT ACT_SITE 149 149
FT ACT_SITE 350 350
FT ACT_SITE 403 403
FT DISULFID 56 313
FT DISULFID 217 229
FT DISULFID 253 281
FT CARBOHYD 107 107

-----
SERINE CARBOXYPEPTIDASE II-2, CHAIN A.
LINKER PEPTIDE (BY SIMILARITY).
SERINE CARBOXYPEPTIDASE II-2, CHAIN B.
BY SIMILARITY.
BY SIMILARITY.
INTERCHAIN (BY SIMILARITY).
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC... ) (POTENTIAL).

Alignment Scores:
Pred. No.: 1
Score: 256
Percent Similarity: 100.00%
Best Local Similarity: 100.00%

SQ SEQUENCE 436 AA; 48952 MW; E0F82D97E0C34DC9 CRC64;
Alignment Scores:
Pred. No.: 29
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.44%
Indels: 0
Gaps: 0

US-09-712-338-1 (1-1668) x CP22_HORVU (1-436)
QY 298 TTGTGTTGAATGTCGCCCTGGA 321
  |||||
Db 48 LeuTriPLeuAsnGlyGlyProGly 55

RESULT 8
YSS2 CAEEL
ID_YSS2 CAEEL STANDARD; PRT; 470 AA.
AC Q09921;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative serine carboxypeptidase K10B2.2 precursor (EC 3.4.16.-).
GN K10B2.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Miller N.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U28730; AAA68259.1; -.
DR PIR; T16606; T16606.
DR HSSP; P10619; 1IIV.
DR MEROPS; S10.002; -.
DR WormPep; K10B2.2; CE02009.
DR InterPro; IPR000379; Ser esters site.
DR InterPro; IPR001563; Serine carboxpept.
DR Pfam; PF00450; serine carboxpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxpept; 2.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hypothetical protein; Hydrolase; Carboxypeptidase; Glycoprotein;
  Signal.
FT SIGNAL 1 19
FT CHAIN 20 470
FT ACT_SITE 169 169
FT ACT_SITE 380 380
FT ACT_SITE 441 441
FT CARBOHYD 132 132
FT CARBOHYD 316 316
FT CARBOHYD 396 396
FT SEQUENCE 470 AA; 53158 MW; CCC2DACB75EF30FC CRC64;

-----
POTENTIAL.
PUTATIVE SERINE CARBOXYPEPTIDASE K10B2.2.
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC... ) (POTENTIAL).
N-LINKED (GLCNAC... ) (POTENTIAL).
N-LINKED (GLCNAC... ) (POTENTIAL).

Alignment Scores:
Pred. No.: 28.6
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Mismatches: 0
Conservative: 0
Mismatches: 0
```

Query Match: 1.44% Indels: 0
DB: Gaps: 0

US-09-712-338-1 (1-1668) x YSS2_CAEEL (1-470)

QY 298 TTGTGGTTGAATGGTGGCCCTGGA 321

Db 71 LeuTrpLeuAsgGlyProGly 78

RESULT 9

ID PRTP_MOUSE

AC P16675; Q8VRF6; STANDARD; PRT; 474 AA.

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE DE Lysosomal protective protein precursor (EC 3.4.16.5) (Cathepsin A)

GN PGGB. (Carboxypeptidase C) (protective protein for beta-galactosidase).

OS Mus musculus (Mouse)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90170982; PubMed=2106523;

RA Galjart N.J., Gilleman N., Meijer D., D'Azzo A.;

RT "Mouse 'protective protein'. cDNA cloning, sequence comparison, and

RT expression.";

RL J. Biol. Chem. 265:4678-4684 (1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Carninci P., Prange C.,

RA Richards S., Worley K.C., Hale S., Garcia A.L., Gay L.J., Huliy S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

CC -!- FUNCTION: PROTECTIVE PROTEIN APPEARS TO BE ESSENTIAL FOR BOTH THE

CC ACTIVITY OF BETA-GALACTOSIDASE AND NEURAMINIDASE, IT ASSOCIATES

CC WITH THESE ENZYMS AND EXERTS A PROTECTIVE FUNCTION NECESSARY FOR

CC THEIR STABILITY AND ACTIVITY. THIS PROTEIN IS ALSO A

CC CARBOXYPEPTIDASE AND CAN DEAMIDATE TACHYKININS.

CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a

CC broad specificity.

CC -!- SUBUNIT: HETERODIMER OF A 32 kDa AND 20 kDa CHAINS HELD TOGETHER

CC BY DISULFIDE BONDS.

CC -!- SUBCELLULAR LOCATION: Lysosomal.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; J05261; AAA39982.1; -.
DR EMBL; BC018534; AAH18534.1; -.
DR PIR; A35732; A35732.
DR HSP; P10619; LIVY.
DR MEROPS; S10.002; -.
DR MGD; MGI:97748; Ppqb.
DR InterPro; IPR000379; Ser esters site.
DR InterPro; IPR001563; Serine_carbpept.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carbpept; 2.
DR PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal; Lysosome.
FT SIGNAL 1 23
FT CHAIN 24 474 LYSOSOMAL PROTECTIVE PROTEIN.
FT CHAIN 24 320 32 kDa CHAIN.
FT CHAIN 321 474 20 kDa CHAIN.
FT ACT_SITE 173 173 BY SIMILARITY.
FT ACT_SITE 394 394 BY SIMILARITY.
FT ACT_SITE 451 451 BY SIMILARITY.
FT DISULFID 83 356 BY SIMILARITY.
FT DISULFID 235 251 BY SIMILARITY.
FT DISULFID 236 241 BY SIMILARITY.
FT DISULFID 276 325 BY SIMILARITY.
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 425 425 G -> W (IN REF. 2) (POTENTIAL).
SQ SEQUENCE 474 AA; 53844 MW; 8F4D5A7F4FEB6C6E CRC64;

Alignment Scores:
Pred. No.: 28.6 Length: 474
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: Gaps: 0

US-09-712-338-1 (1-1668) x PRTP_MOUSE (1-474)

QY 298 TTGTGGTTGAATGGTGGCCCTGGA 321

Db 75 LeuTrpLeuAsgGlyProGly 82

RESULT 10

ID CBP2_HORVU

AC P08618; P93177; STANDARD; PRT; 476 AA.

DT 01-NOV-1988 (Rel. 09, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serine carboxypeptidase II precursor (EC 3.4.16.6) (Carboxypeptidase

DE D) (CP-MII) [Contains: Serine carboxypeptidase II chain A; Serine

DE carboxypeptidase II chain B].

GN CBP2 OR CXP.2

OS Hordeum vulgare (Barley).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;

OC Triticeae; Hordeum.

OX NCBI_TaxID=4513;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Igri; TISSUE=Etisolated leaf;

RA Roher A., Lok F., Cameron-Mills V., von Wettstein D.;

RL "The gene family of serine carboxypeptidases in barley";

RL Submitted (NOV-1996) to the EMBL/GenBank/DDJ databases.

RN [2]

RP SEQUENCE OF 35-294 AND 314-472.

RA Soerensen S.B., Svendsen I., Breddam K.;

RT "Primary structure of carboxypeptidase II from malted barley.";

RL Carlsberg Res. Commun. 52:285-295 (1987).

CC -!- FUNCTION: MAY BE INVOLVED IN THE DEGRADATION OF SMALL PEPTIDES (2-

CC 5 RESIDUES) OR IN THE DEGRADATION OF STORAGE PROTEINS IN THE

CC

CC CC EMBRYO.

CC CC -1- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine

CC CC or lysine residue.

CC CC -1- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS

CC CC COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND.

CC CC -1- SUBCELLULAR LOCATION: SECRETED INTO THE ENDOSPERM.

CC CC -1- DEVELOPMENTAL STAGE: SIMULTANEOUSLY PRESENT IN ALEURONE AND

CC CC ENDOSPERM BETWEEN 20 AND 30 DAYS POSTANTHESIS. ACCUMULATES IN THE

CC CC DEVELOPING GRAIN AND IS STORED IN ITS ACTIVE FORM IN THE MATURE

CC CC GRAIN. ALSO FOUND IN THE ROOTS AND SHOOTS OF THE GROWING SEEDLING.

CC CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

CC CC -----

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC CC the European Bioinformatics Institute. There are no restrictions on its

CC CC use by non-profit institutions as long as its content is in no way

CC CC modified and this statement is not removed. Usage by and for commercial

CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC CC or send an email to license@isb-sib.ch).

CC CC -----

DR EMBL; Y09602; CAA70815.1; -

DR PIR; T05701; T05701.

DR HSSP; P08819; 1WHT.

DR MEROPS; S10.005; -

DR InterPro; IPR001563; Serine_carbpept.

DR Pfam; PF00450; serine_carbpept; 1.

DR PRINTS; PR00724; CRBOXYPTASEC.

DR ProDom; PD001189; Serine_carbpept; 1.

DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.

DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.

KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal;

KW Multigene family.

FT SIGNAL 1 34

FT CHAIN 35 294

FT PROPEP 295 313 SERINE CARBOXYPEPTIDASE II, CHAIN A.

FT CHAIN 314 476 LINKER PEPTIDE (BY SIMILARITY).

FT ACT_SITE 190 190 SERINE CARBOXYPEPTIDASE II, CHAIN B.

FT ACT_SITE 390 390 BY SIMILARITY.

FT ACT_SITE 443 443 BY SIMILARITY.

FT DISULFID 97 333 INTERCHAIN (BY SIMILARITY).

FT DISULFID 254 266 BY SIMILARITY.

FT CARBOHYD 148 148 BY SIMILARITY.

FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .).

FT CARBOHYD 231 291 N-LINKED (GLCNAC. . .).

FT CARBOHYD 341 341 N-LINKED (GLCNAC. . .).

FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .).

FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (PARTIAL).

FT CARBOHYD 472 472 N-LINKED (GLCNAC. . .).

FT MOD_RES 314 314 BLOCKED.

FT VARIANT 351 352

FT CONFLICT 181 181 TN -> AT.

FT CONFLICT 181 181 Y -> R (IN REF. 2).

SO SEQUENCE 476 AA; 52625 MW; 18685725B1A6B5E4 CRC64;

Alignment Scores:

Pred. No.:	28.6	Length:	476
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.44%	Indels:	0
DB:	1	Gaps:	0

US-09-712-338-1 (1-1668) x CBP2_HORVU (1-476)

Qy 298 TTGTGGTGAATGGTGGCCCTGGA 321

Db 89 LeuTrpLeuAsnGlyGlyProGly 96

|||||

RESULT 11

PRTP_HUMAN

ID PRTP_HUMAN STANDARD; PRT; 480 AA.

AC P10619; Q96KJ2; Q9BW68;

DT 01-JUL-1989 (Rel. 11, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Lysoosomal protective protein precursor (EC 3.4.16.5) (Cathepsin A)

DE (Carboxypeptidase C) (protective protein for beta-galactosidase).

GN PCGB.

OS Homo sapiens (Human).

OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE=88311078; PubMed=3136930;

RA Galjart N.J., Gillemans N., Harris A., van de Horst G.T.J.,

RA Verheijen F.W., Galjaard H., D'Azio A.;

RT "Expression of cDNA encoding the human 'protective protein'

RT associated with lysosomal beta-galactosidase and neuraminidase:

RT homology to yeast proteases.";

RL Cell 54:755-764(1988).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,

RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,

RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,

RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,

RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,

RA Coulson A., Coville G.J., Collier R.E., Connor R.E., Corby N.R.,

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,

RA Ghardam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,

RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,

RA Huckle B., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,

RA Kuy M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,

RA Lehaeslatho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,

RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,

RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,

RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,

RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,

RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,

RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,

RA Tracey A., Tomans A.C., Vaudin M., Wall M., Wallis J.M.,

RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,

RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,

RA Rogers J.;

RT "The DNA sequence and comparative analysis of human chromosome 20.";

RL Nature 414:865-871(2001).

RN [3]

RP SEQUENCE FROM N.A.

RX TISSUE=Skin;

RP MEDLINE=22386257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Sanchez A.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [4]

RP SEQUENCE OF 29-53 AND 327-351.
 RC TISSUE-Plasminogen activator.
 RX MEDLINE=90293074; PubMed=1694176;
 RA Jackman H.L., Tan F., Tamei H., Beurling-Harbury C., Li X.-Y.,
 RA Skidgel R.A., Erdoes E.G.;
 RT "A peptidase in human platelets that deamidates tachykinins. Probable
 RT identity with the lysosomal 'protective protein'.";
 RL J. Biol. Chem. 265:11265-11272(1990).
 [5]
 RN FUNCTION, AND MUTAGENESIS.
 RP MEDLINE=91317848; PubMed=1907282;
 RA Galjart N.J., Morreau H., Willemsen R., Gillemans N., Bonten E.J.,
 RA D'Azzo A.;
 RT "Human lysosomal protective protein has cathepsin A-like activity
 RT distinct from its protective function.";
 RL J. Biol. Chem. 266:14754-14762(1991).
 [6]
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RP MEDLINE=96164441; PubMed=8591035;
 RA Rudenko G., Bonten E., D'Azzo A., Hol W.G.J.;
 RT "Three-dimensional structure of the human 'protective protein':
 RT structure of the precursor form suggests a complex activation
 RT mechanism.";
 RL Structure 3:1249-1259(1995).
 [7]
 RN VARIANT GALACTOSIALIDOSIS VAL-440.
 RP MEDLINE=92097522; PubMed=1756715;
 RX Zhou X.-Y., Galjart N.J., Willemsen R., Gillemans N., Galjaard H.,
 RA D'Azzo A.;
 RT "A mutation in a mild form of galactosialidosis impairs dimerization
 RT of the protective protein and renders it unstable.";
 RL EMO J. 10:4041-4048(1991).
 [8]
 RN VARIANTS GALACTOSIALIDOSIS ARG-49; ARG-65; LEU-90; ASN-249 AND
 RP CYS-395.
 RX MEDLINE=93293970; PubMed=8514852;
 RA Shimoto M., Fukuhara Y., Itoh K., Oshima A., Sakuraba H., Suzuki Y.;
 RT "Protective protein gene mutations in galactosialidosis.";
 RL J. Clin. Invest. 91:2393-2398(1993).
 CC -!- FUNCTION: PROTECTIVE PROTEIN APPEARS TO BE ESSENTIAL FOR BOTH THE
 CC ACTIVITY OF BETA-GALACTOSIDASE AND NEURAMINIDASE. IT ASSOCIATES
 CC WITH THESE ENZYMES AND EXERTS A PROTECTIVE FUNCTION NECESSARY FOR
 CC THEIR STABILITY AND ACTIVITY. THIS PROTEIN IS ALSO A
 CC CARBOXYPEPTIDASE AND CAN DEAMIDATE TACHYKININS.
 CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
 CC broad specificity.
 CC -!- SUBUNIT: HETERODIMER OF A 32 kDa AND 20 kDa CHAINS HELD TOGETHER
 CC BY DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- DISEASE: Defects in PPGB are the cause of galactosialidosis
 CC [MIM:256540], an autosomal recessive disease.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M22960; AAA36476.1; -;
 CC EMBL; AL008726; CRA15501.1; -;
 CC EMBL; BC000597; AAH00597.1; -;
 CC PIR; A31589; A31589;
 CC PDB; 1IVY; 2I-APR-97.
 CC MEROPS; S10.002; -;
 CC GENE; HGNC:9251; PPGB.
 CC MIM; 256540; -;
 CC GO; GO:0005783; C:endoplasmic reticulum; TAS.
 CC GO; GO:0004180; F:carboxypeptidase activity; TAS.
 CC GO; GO:0008047; F:enzyme activator activity; TAS.
 CC GO; GO:0005478; F:intracellular transporter activity; TAS.

DR GO; GO:0006886; P:intracellular protein transport; TAS.
 DR InterPro; IPR000379; Ser_estrs_site.
 DR InterPro; IPR001563; Serine_carbpept.
 DR Pfam; PF00450; serine_carbpept; 1.
 DR PRINTS; PR00724; CRBOXYPTASEC.
 DR ProDom; PD001189; Serine_carbpept; 2.
 DR PROSITE; PS001131; CARBOXYPEPT_SER_SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
 DR KX Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal; Lysosome;
 KW 3D-structure; Disease mutation.
 FT SIGNAL 1 28
 FT CHAIN 29 480 LYSOSOMAL PROTECTIVE PROTEIN.
 FT CHAIN 29 326 32 kDa CHAIN.
 FT CHAIN 327 480 20 kDa CHAIN.
 FT ACT SITE 178 178
 FT ACT SITE 400 400 BY SIMILARITY.
 FT ACT SITE 457 457
 FT ACT SITE 457 457
 FT DISULFID 88 362
 FT DISULFID 240 256
 FT DISULFID 241 246
 FT DISULFID 281 331
 FT CARBOHYD 145 145
 FT CARBOHYD 333 333
 FT VARIANT 49 49
 FT VARIANT 65 65
 FT VARIANT 90 90
 FT VARIANT 249 249
 FT VARIANT 395 395
 FT VARIANT 440 440
 FT MUTAGEN 178 178
 FT MUTAGEN 457 457
 FT CONFLICT 19 19
 FT CONFLICT 56 56
 FT HELIX 31 33
 FT STRAND 34 34
 FT TURN 39 40
 FT STRAND 49 55
 FT TURN 58 59
 FT STRAND 60 67
 FT TURN 73 75
 FT STRAND 78 82
 FT TURN 85 87
 FT STRAND 88 88

Alignment Scores:
 Pred. No.: 28.5 Length: 480
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.44% Indels: 0
 DB: 1 Gaps: 0

US-09-712-338-1 (1-1668) x PRTP_HUMAN (1-480)
 QY 298 TTCTGTTGATGCTGCGCTGGA 321
 DB 80 LeutripLeuAsnGlyGlyProGly 87
 RESULT 12
 CBP1_HORVU
 ID -CBP1_HORVU STANDARD; PRT; 499 AA.
 AC P07539; P07520;
 DT 01-APR-1998 (Rel. 07, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)

Serine carboxypeptidase I precursor (EC 3.4.16.5) (Carboxypeptidase C) (CP-MI).
 CBPI OR CAP;1.
 Hordeum vulgare (Barley).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Hordeum.
 NCBI_TaxID=4513;
 [1]
 SEQUENCE FROM N.A.
 TISSUE=Aleurone;
 Rocher A., Lok F., Cameron-Mills V., von Wettstein D.;
 Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 [2]
 SEQUENCE OF 88-499 FROM N.A.
 MEDLINE=88298749; PubMed=3403516;
 Doan N.P., Fincher G.B.;
 "The A- and B-chains of carboxypeptidase I from germinated barley originate from a single precursor polypeptide.";
 J. Biol. Chem. 263:11106-11110(1988).
 [3]
 SEQUENCE OF 31-296 AND 352-499.
 Scerensen S.B., Breddam K., Svendsen I.;
 "Primary structure of carboxypeptidase I from malted barley.";
 Carlsberg Res. Commun. 51:475-485(1986).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE DEGRADATION OF SMALL PEPTIDES (2-5 RESIDUES) OR IN THE DEGRADATION OF STORAGE PROTEINS IN THE EMBRYO.
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a broad specificity.
 CC -1- SUBUNIT: CARBOXYPEPTIDASE I IS A DIMER, WHERE EACH MONOMER IS COMPOSED OF TWO CHAINS LINKED BY DISULFIDE BONDS.
 CC -1- SUBCELLULAR LOCATION: SECRETED INTO THE ENDOSPERM.
 CC -1- DEVELOPMENTAL STAGE: AFTER ONE DAY OF GERMINATION, MAINLY FOUND IN THE SCUTELLUM OF THE DEVELOPING GRAIN; BARELY DETECTABLE AFTER FOUR DAYS, AND ABSENT FROM THE MATURE GRAIN. A LOWER LEVEL OF EXPRESSION IS SEEN IN THE ALEURONE BOTH DURING DEVELOPMENT AND GERMINATION.
 CC -1- PTM: THREE DISULFIDE BONDS ARE PRESENT.
 CC -1- PTM: THE LINKER PEPTIDE IS ENDOPROTEOLYTICALLY EXCISED DURING ENZYME MATURATION.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 EMBL; Y09603; CAA70816.1; -;
 EMBL; J03897; AAA32940.1; -;
 PIR; T05367; CPBS.
 DR HSSP; P08819; 1WHT.
 DR MEROPS; S10.004; -;
 DR InterPro; IPR000379; Ser esters site.
 DR InterPro; IPR001563; Serine carboxpept.
 DR Pfam; PF00450; serine carboxpept; 1.
 DR PRINTS; PR00724; CRBOXPTASEC.
 DR ProDom; PD001189; Serine carboxpept; 2.
 DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
 DR Hydrolase; Carboxypeptidase; Zymogen; Signal.
 FT SIGNAL 1 30
 FT CHAIN 31 296 SERINE CARBOXYPEPTIDASE I, CHAIN A.
 FT PROPEP 297 351 LINKER PEPTIDE.
 FT CHAIN 352 499 SERINE CARBOXYPEPTIDASE I, CHAIN B.
 FT ACT_SITE 188 188 BY SIMILARITY.
 FT ACT_SITE 423 423 BY SIMILARITY.
 FT ACT_SITE 476 476 BY SIMILARITY.
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .).

FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .).
 FT SITE 497 499 MICROBODY TARGETING SIGNAL (POTENTIAL).
 FT CONFLICT 102 102 H -> P (IN REF. 3).
 SQ SEQUENCE 499 AA; 54096 MW; 9C6674B14D9DB9BF CRC64;
 Alignment Scores:
 Pred. No.: 28.3 Length: 499
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.44% Indels: 0
 DB: 1 Gaps: 0
 US-09-712-338-1 (1-1668) x CBPI_HORVU (1-499)
 Qy 298 TTGTGGTTGAATGGTGCCTGGA 321
 Db 84 LeuTriLeuAsnGlyGlyProGly 91
 RESULT 13
 CBPI_ORYSA
 ID CBPI_ORYSA STANDARD; PRT; 510 AA.
 AC P37890;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Serine carboxypeptidase I precursor (EC 3.4.16.5) (Carboxypeptidase C).
 DE CBPI.
 GN Oryza sativa (Rice).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Yukihikari;
 RX MEDLINE=94213891; PubMed=8161571;
 RA Washio K., Ishikawa K.;
 RT "Cloning and sequencing of the gene for type I carboxypeptidase in rice.";
 RL Biochim. Biophys. Acta 1199:311-314(1994).
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a broad specificity.
 CC -1- PTM: THREE DISULFIDE BONDS ARE PRESENT (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 EMBL; D17586; BAA04510.1; -;
 PIR; S43516; S43516.
 DR HSSP; P08819; 1WHT.
 DR MEROPS; S10.004; -;
 DR Gramene; P37890; -;
 DR InterPro; IPR000379; Ser esters site.
 DR InterPro; IPR001563; Serine carboxpept.
 DR Pfam; PF00450; serine carboxpept; 1.
 DR PRINTS; PR00724; CRBOXPTASEC.
 DR ProDom; PD001189; Serine carboxpept; 2.
 DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
 DR Hydrolase; Carboxypeptidase; Zymogen; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 36 POTENTIAL.
 FT PROPEP 37 510 SERINE CARBOXYPEPTIDASE I.
 FT ACT_SITE 194 194 BY SIMILARITY.
 FT ACT_SITE 434 434 BY SIMILARITY.

FT ACT SITE 487 487 BY SIMILARITY. 510
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 508 510 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 510 AA; 55709 MW; 28896247FA1371CF CRC64;

Alignment Scores: 28.2 Length: 510
Pred. No.: 8.00 Matches: 8
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 1.44% Indels: 0
Query Match: 1 Gaps: 0
DB: 1

US-09-712-338-1 (1-1668) x CP23_ORVSA (1-510)
QY 298 TTGTGGTTGAATGGTGGCCCTGGA 321
DB 90 LeuTrpLeuAsnGlyGlyProGly 97

RESULT 14
CP23 HORVU
ID CP23 HORVU STANDARD; PRT; 516 AA.
AC P52711;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase II-3 precursor (EC 3.4.16.6) (CP-MII.3).
GN EXP.2-3.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Alexis; TISSUE=Grain;
RX MEDLINE=94336715; PubMed=7520177;
RA Dal Degan F., Rother A., Cameron-Mills V., von Wettstein D.;
RT "The expression of serine carboxypeptidases during maturation and
germination of the barley grain."
RL Proc. Natl. Acad. Sci. U.S.A. 91:8209-8213(1994).
CC -!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine
or lysine residue.
CC -!- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS
COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE GERMINATING EMBRYO. ALSO
FOUND IN THE ROOTS AND SHOOTS OF THE GROWING SEEDLING.
CC -!- PTM: THE LINKER PEPTIDE IS ENDOPROTEOLYTICALLY EXCISED DURING
ENZYME MATURATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: X78877; CAA55478.1; --
PIR: S44191; S44191.
DR HSP; P08819; 1WHT.
DR MEROPS; S10.005; --
DR InterPro; IPR000379; Ser esters site.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXYPRTASEC.
DR ProDom; PD001189; Serine carbpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal;

KW Multigene family. 20
FT SIGNAL 1 1 POTENTIAL.
FT PROPEP 21 77 POTENTIAL.
FT CHAIN 78 341 SERINE CARBOXYPEPTIDASE II-3, CHAIN A.
FT PROPEP 342 352 LINKER PEPTIDE (BY SIMILARITY). CHAIN B.
FT CHAIN 353 516 SERINE CARBOXYPEPTIDASE II-3, CHAIN B.
FT ACT SITE 236 236 BY SIMILARITY.
FT ACT_SITE 427 427 BY SIMILARITY.
FT ACT_SITE 484 484 BY SIMILARITY.
FT DISULFID 143 391 INTERCHAIN (BY SIMILARITY).
FT DISULFID 300 315 BY SIMILARITY.
FT DISULFID 339 359 BY SIMILARITY.
FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 516 AA; 55913 MW; D41AA1C56CF8D355 CRC64;

Alignment Scores: 28.2 Length: 516
Pred. No.: 8.00 Matches: 8
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 1.44% Indels: 0
Query Match: 1 Gaps: 0
DB: 1

US-09-712-338-1 (1-1668) x CP23_HORVU (1-516)
QY 298 TTGTGGTTGAATGGTGGCCCTGGA 321
DB 135 LeuTrpLeuAsnGlyGlyProGly 142

RESULT 15
CBPY PICPA
ID CBPY PICPA STANDARD; PRT; 523 AA.
AC P52710;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carboxypeptidase Y precursor (EC 3.4.16.5) (Carboxypeptidase YSCV).
GN PRC1.
OS Pichia pastoris (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4922;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 108-142.
RC STRAIN=GTS115;
RX MEDLINE=96381245; PubMed=8789258;
RA Ohi H., Ohtani W., Okazaki N., Furuhashi T., Ohmura T.;
RT "Cloning and characterization of the Pichia pastoris PRC1 gene
encoding carboxypeptidase Y."
RL Yeast 12:31-40(1996).
CC -!- FUNCTION: INVOLVED IN DEGRADATION OF SMALL PEPTIDES.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
broad specificity.
CC -!- SUBCELLULAR LOCATION: LYSOSOME-LIKE VACUOLES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: X87987; CAA61240.1; --
PIR: S61713; S61713.
DR HSP; P00729; 1YSC.
DR MEROPS; S10.001; --
DR InterPro; IPR000379; Ser esters site.
DR InterPro; IPR001563; Carboxypeptidase Y.

Search completed: November 21, 2003, 17:51:51
Job time : 25 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:47:05 ; Search time 31.5 Seconds
(without alignments)
10184.721 Million cell updates/sec

Title: US-09-712-338-1
Perfect score: 555
Sequence: 1 atcgctggtagcaattctt.....ccagtggtgtaggcataag 1668

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 9616862 residues

Word size: 1
Total number of hits satisfying chosen parameters: 565704

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q/cgn2.1/USPTO spool.p/US09712338/runat 17112003.170226 10786/app_query.fasta.1.1863
-DB=PIR 76 -OPMT=fastan -SUFFIX=oli.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HSPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09712338 @CGN 1.1.59 @runat 17112003.170226 10786 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR 76.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	1.8	423	2 S38953	carboxypeptidase D
2	10	1.8	460	1 S31516	serine-type carbox
3	9	1.6	405	2 S70150	tniQ protein - Xan
4	9	1.6	405	2 T08521	tniQ protein homol
5	9	1.6	468	2 D84503	probable serine ca
6	9	1.6	479	2 G96501	probable serine ca
7	9	1.6	495	2 S32179	tniQ protein homol
8	9	1.6	508	1 S46008	probable carboxype
9	2	1.6	2105	2 T18968	probable serine-ty
10	8	1.4	36	2 S78239	ycf32 protein - Od
11	8	1.4	91	2 G90064	hypothetical prote
12	8	1.5	116	2 A88450	protein F2H11.1 (
13	8	1.4	166	2 E71040	hypothetical prote
14	8	1.5	184	2 H69106	conserved hypothet

15	8	1.4	210	2 C84404	hypothetical prote
16	8	1.4	284	2 T49055	serine carboxypept
17	8	1.4	270	2 E75259	hypothetical prote
18	8	1.4	281	2 G82201	hypothetical prote
c	19	1.5	290	2 AD0137	succinate-CoA liga
20	8	1.4	292	2 A71065	probable sugar-bin
c	21	1.5	327	2 G83058	probable permealase
22	8	1.4	334	2 T19127	hypothetical prote
23	8	1.4	356	2 I49340	MIP-1 alpha recept
24	8	1.4	423	1 A29639	carboxypeptidase D
25	8	1.4	425	2 F85360	SERINE CARBOXYPEPT
26	8	1.4	425	2 E84631	probable serine ca
27	8	1.4	447	2 G84772	probable serine ca
28	8	1.4	452	2 H84772	probable serine ca
29	8	1.4	456	2 H86406	probable serine ca
30	8	1.4	456	2 A85139	hypothetical prote
31	8	1.4	458	2 F84746	probable serine ca
c	32	1.5	459	2 H97572	probable serine ca
33	8	1.4	465	2 G82444	hypothetical prote
34	8	1.4	465	2 B85358	SERINE CARBOXYPEPT
35	8	1.4	470	2 T16606	probable serine ca
36	8	1.4	470	2 B96637	hypothetical prote
37	8	1.4	474	2 A35732	protective protein
38	8	1.4	476	1 T05701	carboxypeptidase D
39	8	1.4	477	2 D70546	hypothetical prote
40	8	1.4	480	2 A31589	carboxypeptidase C
41	8	1.4	480	2 T50511	serine-type carbox
c	42	1.5	491	2 AI2793	conserved hypothet
43	8	1.4	499	1 CP8HS	carboxypeptidase C
44	8	1.4	501	2 T43081	serine-type carbox
45	8	1.4	502	2 T43180	serin carboxypepti

ALIGNMENTS

RESULT 1

S38953
carboxypeptidase D (EC 3.4.16.6) - Penicillium janthineillum
N/A alternate names: carboxypeptidase S1
C/Species: Penicillium janthineillum
C/Date: 18-Feb-1994 #sequence_revision 19-Apr-1996 #text_change 31-Mar-2000
C/Accession: S38953
R/Svendsen, I.; Hofmann, T.; Endrizzi, J.; Remington, S.J.; Breddam, K.
F8BS Lett. 333, 39-43, 1993
A/Title: The primary structure of carboxypeptidase S1 from Penicillium janthineillum.
A/reference number: S38953; MUID:94039747; PMID:8224168
A/Accession: S38953
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-423 <SVS>
C/Superfamily: serine carboxypeptidase
C/Keywords: hydrolase; serine carboxypeptidase

Alignment Scores:

Pred. No.: 0.356 Length: 423
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x S38953 (1-423)

Qy 679 ACGAGAGCTATGAGGGCACTATGGTCCT 708
|||||
Db 141 ThrGluSerTyrGlyGlyHisTyrGlyPro 150
|||||

RESULT 2

S51516
serine-type carboxypeptidase (EC 3.4.16.-) Z precursor - Absidia zychae
C/Species: Absidia zychae
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Jun-2000
C/Accession: S51516; S78013; S78014

R;Lee, B.R.; Takeuchi, M.; Kobayashi, Y.
 Curr. Genet. 27, 159-165, 1995
 A:Title: Molecular cloning and sequence analysis of the scpZ gene encoding the serine carboxypeptidase from *Xanthomonas* sp.
 A:Reference number: S51516; MUID:95308538; PMID:7783719
 A:Accession: S51516
 A:Molecule type: DNA
 A:Residues: 1-460 <LEE>
 A:Cross-references: EMBL:D16519; NID:9556466; PIDN:BA03966.1; PID:9995456
 A:Experimental source: strain NRIC 1199
 A:Accession: S78013
 A:Molecule type: protein
 A:Residues: 52-62;90-99;367-381 <LEB>
 A:Accession: S78014
 A:Molecule type: mRNA
 A:Residues: 18-460 <LES>
 C:Genetics:
 A:Gene: scpZ
 A:Introns: 64/3; 95/1; 107/2; 138/3; 234/2; 246/1; 270/1; 343/3; 353/2; 384/3; 437/3
 C:Superfamily: serine carboxypeptidase
 C:Keywords: glycoprotein; hydrolase; serine carboxypeptidase
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-51/Domain: propeptide #status predicted <PR>
 F:52-460/Product: serine-type carboxypeptidase #status experimental <MAT>
 F:40,46,119/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:186/Active site: Ser #status predicted

Alignment Scores: 0.351 Length: 460
 Pred. No.: 10.00 Matches: 10
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 1.80% Gaps: 0
 DB: 1

US-09-712-338-1 (1-1668) x S51516 (1-460)

QY 1456 CCGCTTATGAGCGAGGCCATGAAGTCCCA 1485
 |||||
 Db 430 ArgValtyrGluAlaGlyHisGluValPro 439

RESULT 3
 S70150
 tniQ protein - Xanthomonas sp.
 C:Species: Xanthomonas sp.
 C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 11-Jan-2000
 C:Accession: S70150
 R:Kholodii, G.Y.; Mindlin, S.Z.; Bass, I.A.; Yurieva, O.V.; Minakhina, S.V.; Nikiforov, M.; Microbiol. 17, 1189-1200, 1995
 A:Title: Four genes, two ends, and a res region are involved in transposition of Tn5053:
 A:Reference number: S70140; MUID:96130850; PMID:8594337
 A:Accession: S70150
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-405 <KHO>
 A:Cross-references: EMBL:U40585; NID:9710572; PIDN:AA98331.1; PID:9710584
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
 C:Genetics:
 A:Gene: tniQ
 A:Start codon: GTG
 C:Superfamily: Xanthomonas tniQ protein
 C:Keywords: transposition

Alignment Scores: 3.86 Length: 405
 Pred. No.: 9.00 Matches: 9
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 1.62% Gaps: 0
 DB: 2

US-09-712-338-1 (1-1668) x S70150 (1-405)

QY 262 GAAGCCAGACATACCCAGAACTGCA 288
 |||||
 Db 341 GluAlaArgHisAsnProGluThrAla 349

Db 341 GluAlaArgHisAsnProGluThrAla 349
 |||||
 RESULT 4
 T08521
 tniQ protein homolog - Enterobacter aerogenes plasmid R751
 C:Species: Enterobacter aerogenes
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: T08521
 R:Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.
 Plasmid 36, 95-111, 1996
 A:Title: Conservation of the genetic switch between replication and transfer genes of
 A:Reference number: Z16434; MUID:97118926; PMID:8954881
 A:Accession: T08521
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-405 <THO>
 A:Cross-references: EMBL:U67194; NID:91572520; PIDN:AAC64465.1; PID:91572570
 C:Genetics:
 A:Genome: plasmid R751
 C:Superfamily: Xanthomonas tniQ protein
 C:Keywords: transposition

Alignment Scores: 3.86 Length: 405
 Pred. No.: 9.00 Matches: 9
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 1.62% Gaps: 0
 DB: 2

US-09-712-338-1 (1-1668) x T08521 (1-405)

QY 262 GAAGCCAGACATACCCAGAACTGCA 288
 |||||
 Db 341 GluAlaArgHisAsnProGluThrAla 349

RESULT 5
 D84503
 probable serine carboxypeptidase II [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: D84503
 R:Jin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: D84503
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-468 <STO>
 A:Cross-references: GB:AE002093; NID:94733981; PIDN:AAD28662.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g12480
 A:Map position: 2
 C:Superfamily: serine carboxypeptidase

Alignment Scores: 3.77 Length: 468
 Pred. No.: 9.00 Matches: 9
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 1.62% Gaps: 0
 DB: 2

US-09-712-338-1 (1-1668) x D84503 (1-468)

QY 295 ACATTGTGTTGATGTGCGCCCTGCA 321
 |||||
 Db 77 ThrLeuTrpLeuAsnGlyGlyProGly 85

RESULT 6

G96501
 Probable serine carboxypeptidases [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: G96501
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, N.; Anen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: G96501
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-479 <STO>
 A:Cross-references: GB:AE005173; NID:g7523661; PIDN:AAF63101.1; GSPDB:GN00141
 A:Gene: F28H19.5
 A:Map position: 1
 C:Superfamily: serine carboxypeptidase

Alignment Scores:

Pred. No.:	3.76	Length:	479
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.62%	Indels:	0
DB:	2	Gaps:	0

US-09-712-338-1 (1-1668) x G96501 (1-479)

QY 295 ACATGCTGTTGAATGGTGGCCCTGA 321

Db 83 ThrLeuTrpLeuAsnGlyGlyProGly 91

RESULT 7

S32179
 tniQ protein homolog - Klebsiella pneumoniae
 C:Species: Klebsiella pneumoniae
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 11-Jan-2000
 C:Accession: S32179
 R:Radstroem, P.; Sundstrom, L.; Swedberg, G.; Flensburg, J.; Skold, O.
 submitted to the EMBL Data Library, March 1993
 A:Description: The integrons of R751 and Tn21 are transposons related to the retroelement
 A:Reference number: S32177
 A:Accession: S32179
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-495 <RAD>
 A:Cross-references: EMBL:X72585; NID:g288626; PIDN:CAA51177.1; PID:g288629
 C:Superfamily: Xanthomonas tniQ protein
 C:Keywords: transposition

Alignment Scores:

Pred. No.:	3.74	Length:	495
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.62%	Indels:	0
DB:	2	Gaps:	0

US-09-712-338-1 (1-1668) x S32179 (1-495)

QY 262 GAAGCCAGACATAACCCAGAACTGCA 288

Db 431 GiualaArgHisAnProGluThraLa 439

RESULT 8

S46008
 Probable carboxypeptidase C (EC 3.4.16.5) YBR139w - Yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YBR1015; hypothetical protein YBR139w
 C:Species: Saccharomyces cerevisiae
 C:Date: 26-Aug-1994 #sequence_revision 05-Jan-1996 #text_change 19-Apr-2002
 C:Accession: S46008; S46581
 R:Becam, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulski, M.
 submitted to the Protein Sequence Database, August 1994
 A:Reference number: S45995
 A:Accession: S46008
 A:Molecule type: DNA
 A:Residues: 1-508 <BEC>
 A:Cross-references: EMBL:Z36008; NID:g536435; PIDN:CAA85097.1; PID:g536436; GSPDB:GN0000
 A:Experimental source: strain S288C
 R:Becam, A.M.; Cullin, C.; Grzybowska, E.; Lacroute, F.; Nasr, F.; Ozier-Kalogeropoulos, yeast 10(Suppl.A), S1-S11, 1994
 A:Title: The sequence of 29.7kb from the right arm of chromosome II reveals 13 complete
 A:Reference number: S46569; MUID:94378717; PMID:8091856
 A:Accession: S46581
 A:Molecule type: DNA
 A:Residues: 1-508 <BE2>
 A:Cross-references: EMBL:X75891; NID:g456856; PIDN:CAA53497.1; PID:g456869
 A:Experimental source: strain S288C
 C:Genetics:
 A:Cross-references: SGD:S0000343
 A:Map position: 2R
 A:Note: MIPS:YBR139w
 C:Superfamily: serine carboxypeptidase
 C:Keywords: hydrolase; serine carboxypeptidase; transmembrane protein
 F:6-22/Domain: transmembrane #status predicted <TM>
 F:219,415,474/Active site: Ser, Asp, His #status predicted

Alignment Scores:

Pred. No.:	3.73	Length:	508
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.62%	Indels:	0
DB:	1	Gaps:	0

US-09-712-338-1 (1-1668) x S46008 (1-508)

QY 439 CAGCATTTGGAGTCGGTTTCATAT 465

Db 171 GlnProLeuGlyValGlyPheSerTyr 179

RESULT 9

T18968
 probable serine-type carboxypeptidase (EC 3.4.16.-) Y16B4A.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Mar-2000
 C:Accession: T18968; T23145; T26477
 R:Thomas, K.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z19053
 A:Accession: T18968
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-2105 <MIL>
 A:Cross-references: EMBL:Z70203; PIDN:CAA94110.1; GSPDB:GN00028; CESP:Y16B4A.2
 A:Experimental source: clone C05G5
 R:Lloyd, C.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z19697
 A:Accession: T23145
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-2105 <W12>
 A:Cross-references: EMBL:Z93339; PIDN:CAB07544.1; GSPDB:GN00028; CESP:Y16B4A.2
 A:Experimental source: clone H40D05
 R:Wallis, J.
 submitted to the EMBL Data Library, June 1998

Mon Nov 24 13:41:01 2003

A:Reference number: Z20220
A:Accession: T26477
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2105 <WI3>
A:Cross-references: EMBL:AL023825; PIDN:CAA19443.1; GSPDB:GN00028; CESP:Y16B4A.2
A:Experimental source: clone Y16B4A
C:Comment: This protein has 4 repeats of sequence homologous to serine carboxypeptidase.
C:Genetics:
A:Gene: CESP:Y16B4A.2
A:Map position: X
A:Introns: 43/3; 67/2; 179/1; 208/3; 239/2; 281/2; 296/3; 346/2; 444/3; 503/3; 588/3; 681/64/3; 1752/2; 1794/3; 1834/3; 1882/2; 1905/2; 1942/3; 1982/3
C:Keywords: duplication; hydrolase; serine carboxypeptidase

Alignment Scores:
Pred. No.: 3 Length: 2105
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x T18968 (1-2105)
QY 298 TTGTGTTGAATGGTGGCCCTGGAAGC 324
|||||
DB 549 LeuTrpLeuAsnGlyGlyProGlySer 557

RESULT 10
S78239
ycf32 protein - Odontella sinensis chloroplast
C:Species: chloroplast Odontella sinensis
C:Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 20-Jun-2000
C:Accession: S78239; S78293
R:Kowallik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Frieler, U.
Plant Mol. Biol. Rep. 13, 336-342, 1995
A:Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella sinensis
A:Reference number: S78238
A:Accession: S78239
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-36 <KOW>
A:Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91612.1; PID:g1185129
A:Genetics: 5GEN
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
A:Note: 5' gene
A:Accession: S78293
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-36 <KOF>
A:Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91666.1; PID:g1185183
A:Genetics: 3GEN
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
A:Note: 3' gene (c)
A:Genetics: <GEN1>
A:Gene: ycf32
A:Genome: chloroplast
A:Note: gene located on inverted repeat IRA
C:Genetics: <GEN2>
A:Gene: ycf32
A:Genome: chloroplast
A:Note: gene located on inverted repeat IRB
C:Superfamily: hypothetical protein ycf32
C:Keywords: chloroplast

Alignment Scores:
Pred. No.: 60.1 Length: 36
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x S78239 (1-36)
QY 34 TTGTTGCACCCAGTTGGCCCTT 57
|||||
DB 13 LeuValAlaAlaSerTrpAlaLeu 20

RESULT 11
G90064
hypothetical protein [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G90064
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogi-
ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G90064
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-91 <KUR>
A:Cross-references: GB:BA000018; PID:g13702540; PIDN:BA843681.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA2376

Alignment Scores:
Pred. No.: 52.2 Length: 91
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x G90064 (1-91)
QY 767 GGTTCAGCTTAATTTCACTCTC 790
|||||
DB 69 ValPheSerLeuIleSerThrLeu 76

RESULT 12
A88450
protein F21H11.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: A88450
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biol
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.e
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; a
A:Accession: A88450
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-116 <STO>
A:Cross-references: GB:chr III; PIDN:AAA19432.1; PID:g506794; GSPDB:GN00021; CESP:F21H
C:Genetics:
A:Gene: F21H11.1
A:Map position: 3

Alignment Scores:
Pred. No.: 50.3 Length: 116
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x A88450 (1-116)

QY 1151 GCGTCCATGACAGAGTCCTTGCC 1128
|||||
Db 99 AlaSerMetThrGluSerPheAla 106
|||||
RESULT 13
E71040
hypothetical protein PH1613 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C;Accession: E71040
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: E71040
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-166 <KAW>
A;Cross-references: GB:AP000006; NID:G3236133; PIDN:BAA30725.1; PID:G3258042
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1613
Alignment Scores:
Pred. No.: 47.6 Length: 166
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 2 Gaps: 0
US-09-712-338-1 (1-1668) x E71040 (1-166)
QY 1393 CCCTGAAAGTCAACGGCTGCGAG 1416
|||||
Db 43 ProLeuIysValAsnGlyValGlu 50
|||||
RESULT 14
H69106
conserved hypothetical protein MTH1796 - Methanobacterium thermoautotrophicum (strain De
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: H69106
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: H69106
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-184 <MTH>
A;Cross-references: GB:AE000934; GB:AE000666; NID:G2622924; PIDN:AAB86262.1; PID:G262292
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH1796
Alignment Scores:
Pred. No.: 46.8 Length: 184
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 2 Gaps: 0
US-09-712-338-1 (1-1668) x H69106 (1-184)
QY 768 ACCATTACACTACGTTGGCAAT 745
|||||

Db 25 ThrIleAsnThrThrValGlyAsn 32
RESULT 15
C84404
hypothetical protein Vng2543c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: C84404
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jable
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, F.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84404
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-210 <STO>
A;Cross-references: GB:AE004437; NID:G10581934; PIDN:AAG20599.1; GSPDB:GN00138
C;Genetics:
C;Superfamily: Pyrococcus horikoshii hypothetical protein PH0010
Alignment Scores:
Pred. No.: 45.9 Length: 210
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 2 Gaps: 0
US-09-712-338-1 (1-1668) x C84404 (1-210)
QY 162 ACCCGGGCAGAGGGCTGTGGA 185
|||||
Db 57 ThrArgGlyArgGlyArgLeuArg 64
|||||
Search completed: November 21, 2003, 17:55:48
Job time : 37.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:41:14 ; Search time 52 Seconds
(without alignments)
10182.920 Million cell updates/sec

Title: US-09-712-338-1

Perfect score: 555

Sequence: 1 atcggtggtacgaattct.....ccagtgttgatggcatag 1668

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2062474

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2 1/USPTO.spool p/US09712338/runat 17112003.170225 10751/app.query.fasta_1.1863
-DB=A Geneseq 19Jun03 -CFM=fastan -SUFFIX=oli.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.gdi
-LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @CGN 1 1 92 @runat 17112003.170225 10751 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=60
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

A Geneseq 19Jun03: *
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	554	99.8	554	19	AAW56099	A. oryzae ATCC2038
2	35	6.3	35	19	AAW57040	A. oryzae ATCC2038
3	24	4.3	551	23	ABR38864	A. niger serine ca
4	19	3.4	20	19	AAW56100	A. oryzae ATCC2038
5	16	2.9	29	19	AAW57041	A. oryzae ATCC2038
6	10	1.8	526	23	ABR38817	A. niger serine ca
7	9	1.6	144	22	AAU46145	Propionibacterium
8	9	1.6	178	22	AAU46801	Propionibacterium
9	9	1.6	454	21	AGI7034	Arabidopsis thalia
10	9	1.6	461	21	AGI7033	Arabidopsis thalia
11	9	1.6	461	21	AGI7032	Arabidopsis thalia
12	9	1.6	461	21	AGI7031	Arabidopsis thalia
13	9	1.6	473	21	AGI7032	Arabidopsis thalia
14	9	1.6	473	21	AGI7032	Arabidopsis thalia
15	9	1.6	473	21	AGI7032	Arabidopsis thalia
16	9	1.6	508	23	ABR38859	Arabidopsis thalia
17	9	1.6	554	23	ABR38859	Arabidopsis thalia
18	9	1.6	623	23	ABR38865	S. cerevisiae BAX-
19	9	1.6	1513	22	ABR38865	A. niger serine ca
20	8	1.4	69	22	AAU51334	Drosophila melanog
21	8	1.4	78	21	AGI7033	Propionibacterium
22	8	1.4	79	23	ABP00677	Zea mays protein f
23	8	1.5	81	22	AAU45980	Human ORFX protein
24	8	1.4	81	24	ABU07821	Propionibacterium
25	8	1.4	91	22	AAU34418	Corn C-4 methyl st
26	8	1.4	91	22	AAU37071	Staphylococcus aur
27	8	1.4	91	22	AAU37371	Staphylococcus aur
28	8	1.4	91	22	AAU37567	Staphylococcus aur
29	8	1.4	93	22	AAU60975	Human polypeptide
30	8	1.4	102	23	ABG21799	Lung specific prot
31	8	1.4	112	22	ABG21799	Novel human diagno
32	8	1.4	116	23	ABR03424	Human ORFX protein
33	8	1.4	128	21	AGI35382	Zea mays protein f
34	8	1.4	153	22	ABG53530	Human liver peptid
35	8	1.4	153	22	ABR38670	Peptide #6176 enco
36	8	1.4	153	22	ABR23746	Protein #5745 enco
37	8	1.4	153	22	AAW59299	Human brain expres
38	8	1.4	153	22	AAW71846	Human bone marrow
39	8	1.4	153	22	AAW19335	Peptide #5769 enco
40	8	1.4	153	22	AAW32131	Peptide #6168 enco
41	8	1.4	153	23	ABG41659	Human peptide enco
42	8	1.5	210	22	ABG10691	Novel human diagno
43	8	1.4	216	15	AAW72106	Bovine growth horm
44	8	1.4	221	21	AGI33367	Zea mays protein f
45	8	1.4	240	21	AGI6761	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAW56099

ID AAW56099 standard; Protein; 554 AA.

XX AAW56099;

XX AAW56099;

XX 27-AUG-1998 (first entry)

DT						
XX						
DE						A. oryzae ATCC20386 carboxypeptidase I protein.
XX						Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous;
KW						food industry.
XX						Aspergillus oryzae.
OS						
XX						
Key						Location/Qualifiers
FT						1..18
FT						/label= signal

FT Protein 19...554
 FT /label= carboxypeptidase I
 XX XX W09814599-A1.
 XX PD 09-APR-1998.
 XX PF 03-OCT-1997; 97WO-US17977.
 XX PR 27-NOV-1996; 96US-0757534.
 XX PR 04-OCT-1996; 96US-0726880.
 XX XX (NOVO) NOVO NORDISK BIOTECH INC.
 XX PA (NOVO) NOVO-NORDISK AS.
 XX XX Berka R, Blinkovsky A, Brown K, Dammann C, Golightly E;
 XX PI Klotz A, Mathisen TE, Rey M;
 XX XX WPI; 1998-240098/21.
 XX DR N-PSDB; AAV28620.
 XX Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous
 PT substrates, useful for improving flavour of foods
 XX Claim 1; Fig 3; 82pp; English.
 XX This sequence represents carboxypeptidase I from Aspergillus oryzae.
 CC This polypeptide has an optimal activity in the range of pH 3.0-7.5 at
 CC 25 deg. C, optimal activity in the range of 55-60 deg. C at pH 4, and a
 CC residual activity of at least 65.5% after 30 minutes at pH 4.0 and
 CC 60 deg. C. It also has the capacity to hydrolyse X from N-CBZ-Ala-X where
 CC N-CBZ is N-carboxybenzyl and X is any amino acid. The carboxypeptidases
 CC can be used for obtaining hydrolysates (which can be enriched in free
 CC glutamic acid or peptide bound glutamic acid residues) from proteinaceous
 CC substrates. The carboxypeptidases can be used in flavour-improving
 CC compositions in the food industry. The products can also be used for the
 CC production of polypeptides free of carboxypeptidase activity.
 XX SQ Sequence 554 AA;

Alignment Scores:
 Pred. No.: 0 Length: 554
 Score: 554.00 Matches: 554
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.82% Indels: 0
 DB: 19 Gaps: 0

US-09-712-338-1 (1-1668) x AAW56099 (1-554)

QY 1 ATCGGTGGCTAGCAATTTCTCAGTGTACCCCTTGGTTCACGCCAGTTGGGCCCTTCCA 60
 Db 1 MetArgGlyTyrGluPheLeuSerValLeuProLeuValAlaAlaSerTrpAlaLeuPro 20
 QY 61 GGNAGTACACCGCGTCCGTCGTAGACAGACCTACCCAGAACCCACCGGGTCAAG 120
 Db 21 GlySerThrProAlaSerValGlyA:GArgGlnLeuProLysAsnProThrGlyValLys 40
 QY 121 ACTCTTACACCGCAACAAATGTCAACATCCGGTACAAAGAACCCCGGGCGAGGGCGTC 180
 Db 41 ThrLeuThrThrAlaAsnValThrIleArgTyrLysGluProGlyAlaGluGlyVal 60
 QY 181 TGGAGACATACCCGGGTGTCAATCTCTCTGGATATGTGCACACCTCTCCGAGTCC 240
 Db 61 CysGluThrThrProGlyValLysSerTyrSerGlyTyrValAspThrSerProGluSer 80
 QY 241 CATACCTTCTCTGGTTCCTTCGAAGCCAGACATAACCCAGAACTGCACCTATCACATTG 300
 Db 81 HisThrPhePheThrPhePheGluAlaArgHisAsnProGluThrAlaProIleThrLeu 100
 QY 301 TGTTGAATGGTGGCCCTCGAAGCGAATCTTTGATCGGTCTCTTCGAAGAGTTGGGCCT 360
 Db 101 TrpLeuAsnGlyGlyProGlySerAspSerLeuIleGlyLeuPheGluGluLeuGlyPro 120

361 TGCATGTCAATTCGACTTTTGTGATCACTACATCAACCTCACTCGTGGAAAGAGGTCTCC 420
 Db CysHisValAsnSerThrPheAspAspTyrIleAsnProHisSerTrpAsnGluValSer 140
 QY 421 AATTACTATTCTCTCCAGCCATTGGAGTCGGCTTTTTCATATAGTATGATACGGTTGAT 480
 Db AsnLeuLeuPheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAsp 160
 QY 481 GGGTCCATTAAACCTGTAACTGGGTCTGTGAAATTCGAGTTTTCGAGGTTCAGGGC 540
 Db GlySerIleAsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGly 180
 QY 541 CGGTACCCCAACCATGATGCCACCTGTGATCGATATCACTACCAATCTTGGCCGACAGCCGCT 600
 Db ArgTyrProThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaAlaGluAla 200
 QY 601 TGGGAGATCCTGCAAGGATTCCTTAGTGGACTACCTAGCTGGACTCTAGGTGAGTCT 660
 Db TrpGluIleLeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSer 220
 QY 661 AAGGACTTCAGTCTATGACGGAGAGCTATGGAGGGCACTATGGTCTCTGCATCTTCAT 720
 Db LysAspPheSerLeuTrpThrGluSerTyrGlyGlyHisTyrGlyProAlaPheAsn 240
 QY 721 CATTTTACGACGACAGATTCGCAACGGTGTGTAAATGGTGTTCAGCTTAAT 780
 Db HisPheTyrGluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsn 260
 QY 781 TTCAACTCTCTGGGAATATTAAACGGCATCATCGACGAGCGCATCCAGGCCCTTACTAC 840
 Db PheAsnSerLeuGlyIleIleAsnGlyIleIleAspGluAlaIleGlnAlaProTyr 280
 QY 841 CTGTAATTCGTGTGAAACATACCTACGTATCAGGCTGTCAACGAGACCGCTCTACAC 900
 Db ProGluPheAlaValAsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsn 300
 QY 901 TACATCAAGTTTGCACCAACCAATGTCAGGATTTGTCAGGATTTGATTTCCACCTCAAA 960
 Db TyrMetLysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSerThrCysLys 320
 QY 961 CAGACAAACCGCACCGCATTAGCTAGCTAGCTACGCCCTCTGCGCGAAGCCACCAATGTGC 1020
 Db GlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCys 340
 QY 1021 AGGACAATGTTGAGGGGCCATATACGCTTGTGCTGGTGGTGGTGTGTATGATATCGG 1080
 Db ArgAspAsnValGluGlyProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArg 360
 QY 1081 CATCCATATGATGACCCGACTCGCCCAAGTTATTACAAACAAATTTCTGGCAAGGACTCT 1140
 Db HisProTyrAspAspProThrProProSerTyrTyrAsnLysPheLeuAlaLysAspSer 380
 QY 1141 GTCATGAGCGCTATCGCGCTCAACATCAACTACACCCAGTCCCAATATGACGTCTACTAC 1200
 Db ValMetAspAlaIleGlyValAsnIleAsnTyrThrGlnSerAsnAsnAspValTyrTyr 400
 QY 1201 GCTTTCAGCAACACGCGCATTTGTCTGGCCCACTTATTCAGAGACCTCGAGGAGATC 1260
 Db AlaPheGlnGlnThrGlyAspPheValTrpProAsnPheIleGluAspLeuGluIle 420
 QY 1261 CTTGCTCTCCCGTGGTGTCTCCCTCATCTATGGCGACCGCCGATCATCTCAACTGG 1320
 Db LeuAlaLeuProValArgValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTyr 440
 QY 1321 TTCGGCGGTTCAGCCGTTTCCCTCGGTGGCAACTATCCCAAGCCGCCAGTTCGGAAGC 1380
 Db PheGlyGlyGlnAlaValSerLeuAlaAlaAsnTyrSerGlnAlaAlaGlnPheArgSer 460
 QY 1381 GCAGGTGTACAGCCCTCGAAGTCAACGGGTGTGAGTATGGGAAATCTCCGAGTATGCT 1440
 Db AlaGlyTyrThrProLeuLysValAsnGlyValGluTyrGlyGluThrArgGluTyrGly 480

QY 1441 AATTCTCTCTCACTCCGCTGATAGCGAGGCGCATGAAGTCCCATCTACTACCGCCCATC 1500
 Db 481 AsnPhSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProIle 500
 QY 1501 GCCTCCCTGCAATTGTTTAAACCGGACTATCTTCGTTGGGATATCGAGAGGCGCAGAAG 1560
 Db 501 AlaSerLeuGlnLeuPheAsnA-gThrIlePheGlyTyrAspIleAlaGluGlyGlnLys 520
 QY 1561 AGATCTGGCCCGAGCTACAGAGCAATGGAACGCTACAGCTACGCAATACACAGTCTGTC 1620
 Db 521 LysIleTyrProSerTyrLysThrAsnGlyThrAlaThrAlaThrHisThrGlnSerSer 540
 QY 1621 GTGCCCTGCTACCGCTACCAACATGTCACGATGTCAGTGTGGTATG 1662
 Db 541 ValProLeuProThrAlaThrSerMetSerSerValGlyMet 554

RESULT 2

AAW57040

ID AAW57040 standard; Protein; 35 AA.

XX AC AAW57040;

XX DT 27-AUG-1998 (first entry)

DE DE A. oryzae ATCC20386 carboxypeptidase I peptide fragment #3.

XX KW Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous; food industry.

XX OS Aspergillus oryzae.

XX PN WO9814599-Al.

XX PD 09-APR-1998.

XX PF 03-OCT-1997; 97WO-US17977.

XX PR 27-NOV-1996; 96US-0757534.

XX PR 04-OCT-1996; 96US-0726880.

XX PA (NOVO) NOVO NORDISK BIOTECH INC.

XX PA (NOVO) NOVO-NORDISK AS.

XX PI Berka R, Blinkovsky A, Brown K, Dambmann C, Golightly E;

XX PI Klotz A, Mathisen TE, Rey M;

XX DR WPI; 1998-240098/21.

XX PT Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous substrates, useful for improving flavour of foods

XX PS Example 2; Page 42; 82pp: English.

XX CC This sequence represents a peptide fragment from carboxypeptidase I isolated from Aspergillus oryzae and generated from cyanogen bromide cleavage. This polypeptide has an optimal activity in the range of pH 3.0-7.5 at 25 deg. C, optimal activity in the range of 55-60 deg. C at pH 4, and a residual activity of at least 65.5% after 30 minutes at pH 4.0 and 60 deg. C. It also has the capacity to hydrolyse X from N-CBZ-Ala-X where N-CBZ is N-carboxy and X is any amino acid. The carboxypeptidases can be used for obtaining hydrolysates (which can be enriched in free glutamic acid or peptide bound glutamic acid residues) from proteinaceous substrates. The carboxypeptidases can be used in flavour-improving compositions in the food industry. The products can also be used for the production of polypeptides free of carboxypeptidase activity.

XX SQ Sequence 35 AA;

Alignment Scores:

Pred. No.: 1.14e-25 Length: 35

Score: 35.00 Matches: 35

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.31% Indels: 0
 DB: 19 Gaps: 0

US-09-712-338-1 (1-1668) x AAW57040 (1-35)

QY 1147 GAGCTATCGGCTCAACATCACTACACCCAGTCCAATATGAGCTCTACTACGCTTTC 1206
 Db 1 ASPAlaIleGlyValAsnIleAsnTyrThrGlnSerAsnAsnAspValTyrTyrAlaPhe 20

QY 1207 CAGCAACAGGCGACTTGTCTGGCCCACTTCATCGAAGACCTC 1251
 Db 21 GlnGlnThrGlyAspPheValTyrProAsnPheIleGluAspLeu 35

RESULT 3

ABR38864

ID ABR38864 standard; Protein; 551 AA.

XX AC ABR38864;

XX DT 24-APR-2003 (first entry)

DE DE A. niger serine carboxypeptidase polypeptide #3.

XX KW Protease; fungal infection; aspergillosis; food; tanning; detergent;
 KW Protein solubility; viscosity; taste; texture; nutritional value;
 XX EC3.4.16.6.

XX OS Aspergillus niger.

XX PN WO200268623-A2.

XX PD 06-SEP-2002.

XX PF 22-FEB-2002; 2002WO-EP01984.

XX PR 23-FEB-2001; 2001EP-0200657.

XX PR 23-FEB-2001; 2001EP-0200658.

XX PR 23-FEB-2001; 2001EP-0200660.

XX PR 26-FEB-2001; 2001EP-0200706.

XX PR 26-FEB-2001; 2001EP-0200707.

XX PR 26-FEB-2001; 2001EP-0200708.

XX PR 26-FEB-2001; 2001EP-0200719.

XX PR 28-MAR-2001; 2001EP-0000075.

XX PR 28-MAR-2001; 2001EP-0000078.

XX PR 28-MAR-2001; 2001EP-0000080.

XX PR 28-MAR-2001; 2001EP-0000087.

XX PR 28-MAR-2001; 2001EP-0000088.

XX PR 21-MAY-2001; 2001EP-0000156.

XX PR 21-MAY-2001; 2001EP-0000159.

XX PR 21-MAY-2001; 2001EP-0000160.

XX PR 21-MAY-2001; 2001EP-0000162.

XX PR 21-MAY-2001; 2001EP-0000165.

XX PR 21-MAY-2001; 2001EP-0000166.

XX PR 21-MAY-2001; 2001EP-0000168.

XX PR 21-JUN-2001; 2001EP-0000240.

XX PR 21-JUN-2001; 2001EP-0000242.

XX PR 21-JUN-2001; 2001EP-0000244.

XX PR 21-JUN-2001; 2001EP-0000246.

XX PR 12-JUL-2001; 2001EP-0000280.

XX PR 12-JUL-2001; 2001EP-0000285.

XX PR 30-JUL-2001; 2001EP-0000323.

XX PR 30-JUL-2001; 2001EP-0000327.

XX PR 02-AUG-2001; 2001EP-0000341.

XX PR 02-AUG-2001; 2001EP-0000342.

XX PR 02-AUG-2001; 2001EP-0000343.

XX PR 02-AUG-2001; 2001EP-0000344.

XX PR 09-AUG-2001; 2001EP-0000357.

XX PR 16-AUG-2001; 2001EP-0000374.

XX PR 16-AUG-2001; 2001EP-0000377.

XX PR 20-SEP-2001; 2001EP-0000478.

XX PR 20-SEP-2001; 2001EP-0000483.

XX PR 22-OCT-2001; 2001EP-0000552.

FT Modified-site /label= unknown
 FT 10 /label= unknown
 XX
 PN WO9814599-A1.
 XX
 XX 09-APR-1998.
 XX
 XX 03-OCT-1997; 97WO-US17977.
 PF
 XX 27-NOV-1996; 96US-0757534.
 PR
 PR 04-OCT-1996; 96US-0726880.
 XX
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO-NORDISK AS.
 XX
 XX Berka R, Blinkovsky A, Brown K, Dammann C, Golightly E,
 PI Klotz A, Mathisen TE, Rey M;
 XX WFI; 1998-240098/21.
 DR
 XX Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous
 PT substrates, useful for improving flavour of foods
 PT
 XX Example 2; Page 42; 82pp; English.
 PS
 XX This sequence represents a peptide fragment from carboxypeptidase I
 CC isolated from Aspergillus oryzae and generated from cyanogen bromide
 CC cleavage. This polypeptide has an optimal activity in the range of pH
 CC 3.0-7.5 at 25 deg. C optimal activity in the range of 55-60 deg. C at
 CC pH 4, and a residual activity of at least 65.5% after 30 minutes at pH
 CC 4.0 and 60 deg. C. It also has the capacity to hydrolyse X from
 CC N-CBZ-Ala-X where N-CBZ is N-carboxybenzoyl and X is any amino acid. The
 CC carboxypeptidases can be used for obtaining hydrolysates (which can be
 CC enriched in free glutamic acid or peptide bound glutamic acid residues)
 CC from proteinaceous substrates. The carboxypeptidases can be used in
 CC flavour-improving compositions in the food industry. The products can
 CC also be used for the production of polypeptides free of carboxypeptidase
 CC activity.
 XX
 SQ Sequence 29 AA;
 Alignment Scores:
 Pred. No.: Length: 29
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.88% Indels: 0
 DB: 19 Gaps: 0
 US-09-712-338-1 (1-1668) x AAWS7041 (1-29)
 QY 1045 TACGCTTTCGTCGTGTCGTGATGATTCGCATCCATATGAT 1092
 Db 11 TyrAlaPheAlaGlyGlyVallyrAspIleArgHisProtyrAsp 26
 RESULT 6
 ABR38617
 ID ABR38817 standard; Protein; 526 AA.
 XX
 AC ABR38817;
 XX
 XX 24-APR-2003 (first entry)
 DT
 XX
 DE A. niger serine carboxypeptidase polypeptide #1.
 XX
 KW Protease; fungal infection; aspergillosis; food; tanning; detergent;
 KW protein solubility; viscosity; taste; texture; nutritional value;
 KW EC3.4.16.6.
 XX
 OS Aspergillus niger.
 XX
 XX WO200268623-A2.
 PN

XX
 PD 06-SEP-2002.
 XX
 XX 22-FEB-2002; 2002WO-EP01984.
 XX
 XX 23-FEB-2001; 2001EP-0200657.
 PR 23-FEB-2001; 2001EP-0200658.
 PR 23-FEB-2001; 2001EP-0200660.
 PR 26-FEB-2001; 2001EP-0200706.
 PR 26-FEB-2001; 2001EP-0200707.
 PR 26-FEB-2001; 2001EP-0200708.
 PR 26-FEB-2001; 2001EP-0200719.
 PR 28-MAR-2001; 2001EP-0200775.
 PR 28-MAR-2001; 2001EP-0200778.
 PR 28-MAR-2001; 2001EP-0200780.
 PR 28-MAR-2001; 2001EP-0200787.
 PR 28-MAR-2001; 2001EP-0200788.
 PR 21-MAY-2001; 2001EP-0200156.
 PR 21-MAY-2001; 2001EP-0200159.
 PR 21-MAY-2001; 2001EP-0200160.
 PR 21-MAY-2001; 2001EP-0200162.
 PR 21-MAY-2001; 2001EP-0200165.
 PR 21-MAY-2001; 2001EP-0200166.
 PR 21-MAY-2001; 2001EP-0200168.
 PR 21-JUN-2001; 2001EP-0200240.
 PR 21-JUN-2001; 2001EP-0200242.
 PR 21-JUN-2001; 2001EP-0200244.
 PR 12-JUL-2001; 2001EP-0200246.
 PR 12-JUL-2001; 2001EP-0200280.
 PR 30-JUL-2001; 2001EP-0200285.
 PR 30-JUL-2001; 2001EP-0200323.
 PR 02-AUG-2001; 2001EP-0200327.
 PR 02-AUG-2001; 2001EP-0200341.
 PR 02-AUG-2001; 2001EP-0200342.
 PR 02-AUG-2001; 2001EP-0200343.
 PR 02-AUG-2001; 2001EP-0200344.
 PR 09-AUG-2001; 2001EP-0200357.
 PR 16-AUG-2001; 2001EP-0200374.
 PR 16-AUG-2001; 2001EP-0200377.
 PR 20-SEP-2001; 2001EP-0200478.
 PR 20-SEP-2001; 2001EP-0200483.
 PR 22-OCT-2001; 2001EP-0200552.
 PR 22-OCT-2001; 2001EP-0200553.
 PR 22-OCT-2001; 2001EP-0200554.
 PR 22-OCT-2001; 2001EP-0200556.
 PR 22-OCT-2001; 2001EP-0200557.
 PR 15-NOV-2001; 2001EP-0200558.
 PR 21-DEC-2001; 2001EP-0200464.
 PR 21-DEC-2001; 2001EP-0200517.
 XX
 XX (STAM) DSM NV.

Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
 Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
 Maier D, Spreafico F, Folkers U, Hopfer S, Kemmer W, Tan P;
 Stiebler J, Albarg R;

WPI; 2002-723203/78.
 P-PSDB; ABZ78184, ABZ78241.

Novel isolated protease polypeptide useful in laboratory, clinical,
 pharmaceutical, chemical, diagnostic, personal care and industrial
 applications -

Claim 13; Page 271-274; 394pp; English.

The invention relates to a novel isolated protease polypeptide. A
 polypeptide or polynucleotide of the invention is useful for diagnosing a
 fungal infection such as aspergillosis, or as a query sequence to perform
 a search against public databases. A polypeptide of the invention is
 useful in a selected number of industrial or pharmaceutical processes, in
 laboratory or clinical processes, in food industry (baking, brewing,
 cheese manufacturing, meat tenderising), in tanning industry and in the

CC manufacture of biological detergents. A polypeptide may also be useful
CC for improving protein solubility, extraction yields, viscosity or taste,
CC texture, nutritional value, minimising of antigenicity or
CC anti-nutritional factors, colour or functionality as well as processing
CC aspects like filterability of the proteinaceous raw material. The
CC sequences shown in ABR38313-ABR38669 represent the A. niger proteases of
CC the invention.

XX SQ Sequence 526 AA;

Alignment Scores: Length: 526
Pred. No.: 0.808 Matches: 10
Score: 10.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.80% Gaps: 0
DB: 23

US-09-712-338-1 (1-1668) x ABR38817 (1-526)

QY 679 ACGGAGCTATGGAGGCACATGCTCT 708
DB 181 ThrGluSerTyrGlyGlyHisTyrGlyPro 190

RESULT 7
AAU46145
ID AAU46145 standard; Protein; 144 AA.
XX AC AAU46145;
XX DT 27-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein #7041.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX OS Propionibacterium acnes.
XX PN WO200181581-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US12865.
XX PR 21-APR-2000; 2000US-199047P.
XX PR 02-JUN-2000; 2000US-208841P.
XX PR 07-JUL-2000; 2000US-216747P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX DR WPI; 2001-616774/71.
XX DR N-PSDB; AAS59529.
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX PS Example 1; SEQ ID No 7340; 1069pp; English.
XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the

CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 144 AA;

Alignment Scores: Length: 144
Pred. No.: 9.68 Matches: 9
Score: 9.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 1.62% Gaps: 0
DB: 22

US-09-712-338-1 (1-1668) x AAU46145 (1-144)

QY 1224 TGTCTGGCCCACTTCATCGAGACCT 1250
DB 15 CysLeuAlaGlnLeuHisArgArgPro 23

RESULT 8
AAU46801
ID AAU46801 standard; Protein; 178 AA.
XX AC AAU46801;
XX DT 27-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein #7697.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX OS Propionibacterium acnes.
XX PN WO200181581-A2.
XX PD 01-NOV-2001.
XX PF 20-APR-2001; 2001WO-US12865.
XX PR 21-APR-2000; 2000US-199047P.
XX PR 02-JUN-2000; 2000US-208841P.
XX PR 07-JUL-2000; 2000US-216747P.
XX PA (CORI-) CORIXA CORP.
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX DR WPI; 2001-616774/71.
XX DR N-PSDB; AAS59535.
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX PS Example 1; SEQ ID No 7996; 1069pp; English.
XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by

CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertyosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pot_sequences.
 XX

SQ Sequence 178 AA;

Alignment Scores:

Pred. No.:	9.4	Length:	178
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.62%	Indels:	0
DB:	22	Gaps:	0

US-09-712-338-1 (1-1668) x AAU46801 (1-178)

QY 16 TTTCTCTCAGTCTACCTGGTGGCA 42

Db 151 PheLeuSerValLeuProLeuValAla 159

RESULT 9

AAG17034

ID AAG17034 standard; Protein; 454 AA.

AC AAG17034;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 17902.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 21-JUN-1999; 99US-0139763.
 PR 22-JUN-1999; 99US-0139817.
 PR 23-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140354.
 PR 28-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 30-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.

us-09-712-338-1.Oli.rag

Mon Nov 24 13:41:00 2003

```

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145102.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158432.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 8.26 Length: 454
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 21 Gaps: 0

US-09-712-338-1 (1-1668) x AAG17034 (1-454)
QY 295 ACATTGGTGTGAATGTGGCGCTGGA 321
DB 59 ThrLeuTrpLeuAsnGlyGlyProGly 67
|||||
RESULT 10
AAG17033
ID AAG17033 standard; Protein; 461 AA.
XX AAG17033;
AC AAG17033;
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 17901.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX OS Arabidopsis thaliana.
XX EP1033405-A2.
XX PF 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 99US-0132485.
```

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132483.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

```
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 8.24
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.62%
DB: 21
Gaps: 0

US-09-712-338-1 (1-1668) x AAG17033 (1-461)
QY 295 ACATTGGTGGTGAATGGTGGCCCTGGA 321
Db 66 ThrLeuTrpLeuAsnGlyGlyProGly 74

RESULT 11
AAG49759
ID AAG49759 standard; Protein; 461 AA.
XX
AC AAG49759;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62981.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
```

PR 23-JUL-1999; 99US-01451145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148555.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Pred. No.:	8.24	Length:	461
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.62%	Indels:	0
DB:	21	Gaps:	0

US-09-712-338-1 (1-1668) x AAG49759 (1-461)

QY 295 ACATTGCTTGAATGGTGGCCCTGGA 321
Db ||||||| 66 ThrLeuTriLeuAsnGlyProGly 74

RESULT 12

AAG49762
ID AAG49762 standard; Protein; 461 AA.

XX AAG49762;

AC AAG49762;
DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 62985.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; Genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX
EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.
XX

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.

us-09-712-338-1.Oli.rag

Mon Nov 24 13:41:00 2003

PR 14-MAY-1999;	99US-0134218.	PR 23-JUL-1999;	99US-0145224.
PR 14-MAY-1999;	99US-0134219.	PR 26-JUL-1999;	99US-0145276.
PR 14-MAY-1999;	99US-0134221.	PR 27-JUL-1999;	99US-0145913.
PR 14-MAY-1999;	99US-0134370.	PR 27-JUL-1999;	99US-0145918.
PR 18-MAY-1999;	99US-0134768.	PR 28-JUL-1999;	99US-0145919.
PR 19-MAY-1999;	99US-0134941.	PR 02-AUG-1999;	99US-0145951.
PR 20-MAY-1999;	99US-0135124.	PR 02-AUG-1999;	99US-0146386.
PR 21-MAY-1999;	99US-0135353.	PR 02-AUG-1999;	99US-0146388.
PR 24-MAY-1999;	99US-0135629.	PR 02-AUG-1999;	99US-0146389.
PR 25-MAY-1999;	99US-0136021.	PR 03-AUG-1999;	99US-0147038.
PR 27-MAY-1999;	99US-0136382.	PR 04-AUG-1999;	99US-0147204.
PR 28-MAY-1999;	99US-0136788.	PR 04-AUG-1999;	99US-0147302.
PR 01-JUN-1999;	99US-0137222.	PR 05-AUG-1999;	99US-0147192.
PR 03-JUN-1999;	99US-0137528.	PR 05-AUG-1999;	99US-0147260.
PR 04-JUN-1999;	99US-0137508.	PR 06-AUG-1999;	99US-0147303.
PR 07-JUN-1999;	99US-0137724.	PR 06-AUG-1999;	99US-0147416.
PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138540.	PR 09-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 10-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139113.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139452.	PR 12-AUG-1999;	99US-0148341.
PR 16-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148565.
PR 17-JUN-1999;	99US-0139492.	PR 13-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139454.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149223.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139459.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139461.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139462.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139463.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151066.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151080.
PR 21-JUN-1999;	99US-0139817.	PR 30-AUG-1999;	99US-0151303.
PR 22-JUN-1999;	99US-0139899.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140353.	PR 01-SEP-1999;	99US-0151930.
PR 23-JUN-1999;	99US-0140354.	PR 07-SEP-1999;	99US-0152363.
PR 24-JUN-1999;	99US-0140695.	PR 10-SEP-1999;	99US-0153070.
PR 28-JUN-1999;	99US-0140823.	PR 13-SEP-1999;	99US-0153758.
PR 29-JUN-1999;	99US-0140991.	PR 15-SEP-1999;	99US-0154018.
PR 30-JUN-1999;	99US-0141287.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0141842.	PR 20-SEP-1999;	99US-0154779.
PR 01-JUL-1999;	99US-0142154.	PR 22-SEP-1999;	99US-0155139.
PR 02-JUL-1999;	99US-0142055.	PR 23-SEP-1999;	99US-0155486.
PR 06-JUL-1999;	99US-0142390.	PR 24-SEP-1999;	99US-0155659.
PR 08-JUL-1999;	99US-0142803.	PR 28-SEP-1999;	99US-0156458.
PR 09-JUL-1999;	99US-0142920.	PR 29-SEP-1999;	99US-0156596.
PR 12-JUL-1999;	99US-0142977.	PR 04-OCT-1999;	99US-0157117.
PR 13-JUL-1999;	99US-0143542.	PR 05-OCT-1999;	99US-0157753.
PR 14-JUL-1999;	99US-0143624.	PR 06-OCT-1999;	99US-0157865.
PR 15-JUL-1999;	99US-0144005.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144085.	PR 08-OCT-1999;	99US-0158232.
PR 16-JUL-1999;	99US-0144086.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144325.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144684.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0144814.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160767.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145089.	PR 22-OCT-1999;	99US-0160815.
PR 22-JUL-1999;	99US-0145192.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145218.		


```

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156558.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159395.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 8.21 Length: 473
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 21 Gaps: 0

US-09-712-338-1 (1-1668) x AAG17032 (1-473)
QY 295 ACATGTGTTGAATGGTGGCCCTGGA 321
Db 78 ThrLeuTrpLeuAsnGlyGlyProGly 86

RESULT 14
AAG49758
ID AAG49758 standard; Protein; 473 AA.
XX AAG49758;
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62980.
KW protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
```


PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

```
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 8.21 Length: 473
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0
DB: 21 Gaps: 0

US-09-712-338-1 (1-1668) x AAG49758 (1-473)
QY 295 ACATTGCTGATGAATGGTGGCCCTGGA 321
   |||||
DB 78 ThrLeuTrpLeuAsnGlyProGly 86

RESULT 15
AAG49761
ID AAG49761 standard; Protein; 473 AA.
XX AC AAG49761;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 62984.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 06-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143277.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
```

```
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149302.
PR 23-AUG-1999; 99US-0149330.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 8.21      Length: 473
Score: 9.00      Matches: 9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.62%      Indels: 0
DB: 21      Gaps: 0

US-09-712-338-1 (1-1668) x AAG49761 (1-473)
QY 295 ACATTGTGTTGAATGTTGGCCCTGGA 321
   |||||
Db 78 ThrLeuTrpLeuAsnGlyGlyProGly 86

Search completed: November 21, 2003, 17:51:01
Job time : 59 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:49:25 ; Search time 19.5 Seconds
(without alignments)

7238.411 Million cell updates/sec

Title:

Perfect score: 555

Sequence: 1 atgcgtggtaacgaattct.....ccagtgtgtatggcatag 1668

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Word size: 1

Total number of hits satisfying chosen parameters: 557304

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO spool_p/US09712338/runat 17112003 170227 10817/app_query.fasta_1.1863
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=oli.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @CGN 1 1 31 @runat 17112003 170227 10817 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	554	99.8	554	US-08-943-714-2	Sequence 2, Appli
2	35	6.3	35	US-08-943-714-5	Sequence 5, Appli
3	19	3.4	20	US-08-943-714-3	Sequence 3, Appli
4	16	2.9	29	US-08-943-714-6	Sequence 6, Appli
5	10	1.8	423	US-08-943-714-9	Sequence 9, Appli
6	8	1.4	81	US-09-439-554-2	Sequence 2, Appli
7	8	1.4	164	US-09-252-991A-20615	Sequence 20615, A
8	8	1.4	164	US-09-252-991A-30154	Sequence 30154, A
9	8	1.4	283	US-09-107-532A-5418	Sequence 5418, Ap
10	8	1.4	291	US-09-439-554-20	Sequence 20, Appli
11	8	1.5	386	US-09-252-991A-27381	Sequence 27381, A
12	8	1.4	421	US-08-807-263-4	Sequence 4, Appli

13	8	1.4	480	2	US-08-828-488-8	Sequence 8, Appli
14	8	1.4	480	4	US-09-299-689A-8	Sequence 8, Appli
15	8	1.4	480	4	US-09-702-705-336	Sequence 336, App
16	8	1.4	480	4	US-09-736-457-336	Sequence 336, App
17	8	1.4	491	1	US-09-640-305-4	Sequence 4, Appli
18	8	1.4	491	1	US-08-360-673-4	Sequence 4, Appli
19	8	1.4	532	2	US-08-899-324-33	Sequence 33, Appli
20	8	1.4	532	3	US-08-329-892B-33	Sequence 33, Appli
21	8	1.4	557	1	US-08-309-341-2	Sequence 2, Appli
22	8	1.4	557	1	US-08-309-341-4	Sequence 2, Appli
23	8	1.4	557	1	US-08-608-267-2	Sequence 4, Appli
24	8	1.4	557	1	US-08-608-267-4	Sequence 4, Appli
25	8	1.4	557	1	US-08-608-452-2	Sequence 2, Appli
26	8	1.4	557	1	US-08-608-452-4	Sequence 2, Appli
27	8	1.4	557	1	US-08-608-224-2	Sequence 2, Appli
28	8	1.4	557	1	US-08-608-224-4	Sequence 2, Appli
29	8	1.4	557	2	US-08-967-149-2	Sequence 2, Appli
30	8	1.4	557	2	US-08-967-149-4	Sequence 4, Appli
31	8	1.4	654	4	US-09-252-991A-23035	Sequence 23035, A
32	8	1.4	679	4	US-09-252-991A-25400	Sequence 25400, A
33	7	1.3	9	3	US-08-159-339A-528	Sequence 528, App
34	7	1.3	10	3	US-08-159-339A-549	Sequence 549, App
35	7	1.3	19	3	US-08-943-714-4	Sequence 4, Appli
36	7	1.3	34	3	US-08-810-009-24	Sequence 24, Appli
37	7	1.3	35	2	US-08-637-759B-348	Sequence 348, App
38	7	1.3	55	3	US-08-871-355A-348	Sequence 348, App
39	7	1.3	55	4	US-09-201-945-348	Sequence 348, App
40	7	1.3	56	3	US-08-866-545-10	Sequence 10, Appli
41	7	1.3	78	2	US-08-831-575-5	Sequence 5, Appli
42	7	1.3	83	4	US-09-252-991A-19155	Sequence 19155, A
43	7	1.3	85	3	US-08-894-173-54	Sequence 54, Appli
44	7	1.3	85	3	US-08-894-173-55	Sequence 55, Appli
45	7	1.3	85	3	US-08-894-173-56	Sequence 56, Appli

ALIGNMENTS

RESULT 1

US-08-943-714-2
; Sequence 2, Application US/08943714
; Patent No. 6187578

GENERAL INFORMATION:

APPLICANT: Blinkovsky, Alexander

APPLICANT: Berka, Randy

APPLICANT: Rey, Michael

APPLICANT: Golightly, Elizabeth

APPLICANT: Klotz, Alan

APPLICANT: Mathisen, Thomas Erik

APPLICANT: Dambmann, Claus

TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids

TITLE OF INVENTION: Encoding Same

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.

STREET: 405 Lexington Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10174

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/943,714

FILING DATE: 03-OCT-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4990.200-US

TELECOMMUNICATION INFORMATION:

Mon Nov 24 13:41:00 2003

TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 554 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-943-714-2

Alignment Scores:
Pred. No.: 0 Length: 554
Score: 554.00 Matches: 554
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 99.82% Indels: 0
DB: 3 Gaps: 0

US-09-712-338-1 (1-1668) x US-08-943-714-2 (1-554)

QY	1	ATGCGTGGCTACGAATTTCTCTCAGTGTACCTTGGTTCAGCGAGTGGCCCTTCCA	60
DB	1	MetArgGlyTyrGluPheLeuSerValLeuProLeuValAlaAAserTrpAlaLeuPro	20
QY	61	GGAAGTACACCGGCTCGTTCGTAGACAGACAGTACCAAGAACCCACCGGGGTCAAG	120
DB	21	GlySerThrProAlaSerValGlyArgArgGlnLeuProLysAsnProThrGlyVallys	40
QY	121	ACTCTTACACCGCAAAATCTCCATCCGTTCAAGGAACCCGGGGCAGAGGGCTC	180
DB	41	ThrLeuThrThrAlaAsnValThrIleArgTyrLysGluProGlyAlaGluGlyVal	60
QY	181	TCGAGAGTACCCCGGCTGTAATCTCTACTCTGATATGTGCACACCTCTCCGAGTCC	240
DB	61	CysGluThrThrProGlyValLysSerTyrSerGlyTyrValAspThrSerProGluSer	80
QY	241	CATACCTTCTCTGTTCTTCGAAGCCAGACATACCAAGAACTCCACTATCATTG	300
DB	81	HisThrPhePheThrPheGluAlaArgHAsnProGluThrAlaProIleThrLeu	100
QY	301	TGCTTGAATGGTGGCCCTGGAAGCGATTCTTTCATCGGTCTCTTGAAGAGTGGCCCT	360
DB	101	TrpLeuAsnGlyGlyProGlySerAspSerLeuIleGlyLeuPheGluGluGlyPro	120
QY	361	TGCATGTCAATTCGACTTTGTAGTACATCAACCTCCTCGTGGACGAGGTCTCC	420
DB	121	CysHisValAsnSerThrPheAspAspTyrIleAsnProHisSerTrpAsnGluValSer	140
QY	421	AATTTACTATCTCTGCCAGCCATTGGGAGTGGGCTTTTCATATAGTATACGTTGAT	480
DB	141	AsnLeuLeuPheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAsp	160
QY	481	GGGTCCATTAACCTGTACTGGGGTCTCGAATAATTCGAGCTTTTCGAGAGTTCAGGC	540
DB	161	GlySerIleAsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGly	180
QY	541	CGGTACCAACCATTTGATGCCACTCTGATCGATACCAATCTTCCGCGAGAGCGCT	600
DB	181	ArgTyrProThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaAlaGluAlaAla	200
QY	601	TGGAGATCTCTGCAAGATTCTCTAGTGGACTACTAGCTTGGACTTAGGGTGCAGTCT	660
DB	201	TrpGluIleLeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSer	220
QY	661	AAGGACTTCACTCTATCGACGAGAGTATGGAGGCACTATGGTCTCTGCACTTCAAT	720
DB	221	LysAspPheSerLeuTrpThrGluSerTyrGlyGlyHisTyrGlyProAlaPheAsn	240
QY	721	CAITTTTACGAGAGATGAGAGATTCGCAACGGTAGTGTAAATGGTGTTCAGCTTAAT	780

DB	241	HisPheTyrGluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsn	260
QY	781	TTCAACTCTCTGGAAATTTAAGCGCATCATCGACGAGGATCCAGGCCCTTACTAC	840
DB	261	PheAsnSerLeuGlyIleIleAsnGlyIleAspGluAlaIleGlnAlaProTyrTyr	280
QY	841	CCTGAATTCGTGTGAACATACCTACGGTATCAAGCTGTCAACGAGACCGTCTACAAC	900
DB	281	ProGluPheAlaValAsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsn	300
QY	901	TACATGAAGTTTGCACCAACCAATGCCAATGGTTGCCAGGATTTGATTTCCACTGCAAA	960
DB	301	TyrMetLysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSerThrCysLys	320
QY	961	CAGACAAACCGCACCGCATTAGCTAGCTACGCCCTCTCGCGCGAGCCACCAACATGCG	1020
DB	321	GlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCys	340
QY	1021	AGGACCAATGTTGAGGGCCATACAGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1080
DB	341	ArgAspAsnValGluGlyProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArg	360
QY	1081	CATCCATATGATGACCCGACTCCCGCAAGTTATTACAACTTTCTGGGAAAGACTCT	1140
DB	361	HisProTyrAspAspProThrProProSerTyrTyrAsnLysPheLeuAlaLysAspSer	380
QY	1141	GTCATGAGCGTATCGGCTCAACATCACTACACCCAGTCCAAATAAGTACGCTCTACTAC	1200
DB	381	ValMetAspAlaIleGlyValAsnIleAsnTyrThrGlnSerAsnAspValTyrTyr	400
QY	1201	GCTTTCCAGCAAAACGAGCGACTTCTCTGGCCCAACTTCATCGAAGACTCGGAGAGATC	1260
DB	401	AlaPheGlnGlnThrGlyAspPheValTrpProAsnPheIleGluAspLeuGluIle	420
QY	1261	CTTGTCTCTCCCGTCT	1320
DB	421	LeuAlaLeuProValArgValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTrp	440
QY	1321	TTCCGGGTGTCAGGCGGTTTCCCTCTCGGACTACTCCCAAGCCGCCAGTTCGGAAGC	1380
DB	441	PheGlyGlyGlnAlaValSerLeuAlaAlaAsnTyrSerGlnAlaAlaGlnPheArgSer	460
QY	1381	GCAGGTGTCACGCCCTCGAAAGTCAACGGCGTCGAGTATGGGAAACTCCGAGTATGTT	1440
DB	461	AlaGlyTyrThrProLeuLysValAsnGlyValGluTyrGlyGluThrArgGluTyrGly	480
QY	1441	AATTTCT	1500
DB	481	AsnPheSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProIle	500
QY	1501	GCTCTCTCTCAATTTTAAACCGGACTATCTCTGGTTGGATATCGCAGAGGCCAGAG	1560
DB	501	AlaSerLeuGlnLeuPheAsnArgThrIlePheGlyTrpAspIleAlaGluGlyGlnLys	520
QY	1561	AAGATCTGCGCCAGCTACCAAGCAATGGAAACGGCTACAGCTACGATACACAGTCTGTC	1620
DB	521	LysIleTrpProSerTyrLysThrAsnGlyThrAlaThrAlaThrHisThrGlnSerSer	540
QY	1621	GTCGCGCTGCTACCGCTACCGCTACCGCTACCGCTACCGCTACCGCTACCGCTACCGCT	1680
DB	541	ValProLeuProThrAlaThrSerMetSerSerValGlyMet	554

RESULT 2
US-08-943-714-5
; Sequence 5, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berk, Randy
; APPLICANT: Key, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik

CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10174
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/943,714
 FILING DATE: 03-OCT-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 4990.200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEFAX: 212-878-9655
 TELEX:
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: No. 6187578e
 US-08-943-714-3
 Alignment Scores:
 Pred. No.: 2,22e-10 Length: 20
 Score: 19.00 Matches: 19
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.42% Indels: 0
 DB: 3 Gaps: 0
 US-09-712-338-1 (1-1668) x US-08-943-714-3 (1-20)
 QY 688 TATGGAGGCACATATGGTCCTGCATCTTCATCATTTTACGAGCGAATGAGAGA 744
 Db 2 TyrGlyGlyHisTyrGlyProAlaPheAsnHisPheTyrGluGlnAsnGluArg 20
 RESULT 4
 US-08-943-714-6
 Sequence 6, Application US/08943714
 Patent No. 6187578
 GENERAL INFORMATION:
 APPLICANT: Blinkovsky, Alexander
 APPLICANT: Berka, Randy
 APPLICANT: Rey, Michael
 APPLICANT: Golightly, Elizabeth
 APPLICANT: Klotz, Alan
 APPLICANT: Mathisen, Thomas Erik
 APPLICANT: Dambmann, Claus
 TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
 TITLE OF INVENTION: Encoding Same
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
 STREET: 405 Lexington Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10174
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/943,714

FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6187578e
US-08-943-714-6
Alignment Scores:
Pred. No.: 2,27e-07 Length: 29
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.88% Indels: 0
DB: 3 Gaps: 0
US-09-712-338-1 (1-1668) x US-08-943-714-6 (1-29)
QY 1045 TAGCGCTTGTGTCGTCGTGTATGATATTCGGCATCCATATGAT 1092
Db 11 TyrAlaPheAlaGlyArgGlyValTyAspIleArgHisProTyAsp 26
RESULT 5
US-08-943-714-9
Sequence 9, Application US/08943714
Patent No. 6187578
GENERAL INFORMATION:
APPLICANT: Blinkovsky, Alexander
APPLICANT: Berka, Randy
APPLICANT: Key, Michael
APPLICANT: Golightly, Elizabeth
APPLICANT: Klotz, Alan
APPLICANT: Mathisen, Thomas Erik
APPLICANT: Dammann, Claus
TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
TITLE OF INVENTION: Encoding Same
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,714
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-943-714-9
Alignment Scores:
Pred. No.: 0.21 Length: 423
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.80% Indels: 0
DB: 3 Gaps: 0
US-09-712-338-1 (1-1668) x US-08-943-714-9 (1-423)
QY 679 ACGGAGAGCTATGGAGGCACATGTCCT 708
Db 141 ThrGluSerTyTGlyGlyHisTyGlyPro 150
RESULT 6
US-09-439-554-2
Sequence 2, Application US/09439554
Patent No. 6479733
GENERAL INFORMATION:
APPLICANT: Rafalski, Jan Antoni
APPLICANT: Odell, Joan T.
APPLICANT: Sakai, Hajime
APPLICANT: Thorpe, Catherine J.
APPLICANT: Kinney, Anthony J.
APPLICANT: Famodu, Omolayo O.
TITLE OF INVENTION: STEROL METABOLISM ENZYMES
FILE REFERENCE: BB1114 US NA
CURRENT APPLICATION NUMBER: US/09/439,554
CURRENT FILING DATE: 1999-11-12
EARLIER APPLICATION NUMBER: 60/108,351
EARLIER FILING DATE: 1998-No. 6479733ember-13
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 2
LENGTH: 81
TYPE: PRT
ORGANISM: Zea mays
US-09-439-554-2
Alignment Scores:
Pred. No.: 24.2 Length: 81
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 4 Gaps: 0
US-09-712-338-1 (1-1668) x US-09-439-554-2 (1-81)
QY 619 TTCCTTAGTGGACTACTAGCTTG 642
Db 29 PheLeuSerGlyLeuProSerLeu 36
RESULT 7
US-09-252-991A-20615
Sequence 20615, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 20615
 ; LENGTH: 164
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-20615

Alignment Scores:
 Pred. No.: 23.1 Length: 164
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.44% Indels: 0
 DB: 4 Gaps: 0

US-09-712-338-1 (1-1668) x US-09-252-991A-20615 (1-164)

QY 1559 AGAAGATCTGGCCAGCTACAAGA 1582
 |||||
 Db 78 ArgArgSerGlyProAlaThrArg 85

RESULT 8

US-09-252-991A-30154
 ; Sequence 30154, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 30154
 ; LENGTH: 164
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-30154

Alignment Scores:
 Pred. No.: 23.1 Length: 164
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.44% Indels: 0
 DB: 4 Gaps: 0

US-09-712-338-1 (1-1668) x US-09-252-991A-30154 (1-164)

QY 1559 AGAAGATCTGGCCAGCTACAAGA 1582
 |||||
 Db 78 ArgArgSerGlyProAlaThrArg 85

RESULT 9

US-09-107-532A-5418
 ; Sequence 5418, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts

COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD/ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 5418:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 283 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...283
 SEQUENCE DESCRIPTION: SEQ ID NO: 5418:
 US-09-107-532A-5418

Alignment Scores:
 Pred. No.: 22.3 Length: 283
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.44% Indels: 0
 DB: 4 Gaps: 0

US-09-712-338-1 (1-1668) x US-09-107-532A-5418 (1-283)

QY 1135 GACTCTGTCATGGAGCTATGGC 1158
 |||||
 Db 187 AspSerValMetAspAlaIleGly 194

RESULT 10

US-09-439-554-20
 ; Sequence 20, Application US/09439554
 ; Patent No. 6479733
 ; GENERAL INFORMATION:
 ; APPLICANT: Rafalski, Jan Antoni
 ; APPLICANT: Odeil, Joan T.
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Thorpe, Catherine J.
 ; APPLICANT: Kinney, Anthony J.
 ; APPLICANT: Pamodu, Omolayo O.
 ; TITLE OF INVENTION: STEROL METABOLISM ENZYMES
 ; FILE REFERENCE: BB114 US NA
 ; CURRENT APPLICATION NUMBER: US/09/439,554
 ; CURRENT FILING DATE: 1999-11-12
 ; EARLIER APPLICATION NUMBER: 60/108,351
 ; EARLIER FILING DATE: 1998-No. 6479733ember-13
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 20
 ; LENGTH: 291
 ; TYPE: PRT

```

; ORGANISM: Zea mays
US-09-439-554-20

Alignment Scores:
Pred. No.: 22.2 Length: 291
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 4 Gaps: 0

US-09-712-338-1 (1-1668) x US-09-439-554-20 (1-291)
QY 619 TTCCTTAGTGACTACCTAGCTTG 642
DB 61 PheLeuSerGlyLeuProSerLeu 68

RESULT 11
US-09-252-991A-27381
; Sequence 27381, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27381
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27381

Alignment Scores:
Pred. No.: 21.8 Length: 388
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 4 Gaps: 0

US-09-712-338-1 (1-1668) x US-09-252-991A-27381 (1-388)
QY 1587 ATTCGCTTTGTAGCTGGCCAGAT 1564
DB 326 IleArgLeuValAlaGlyProasp 333

RESULT 12
US-08-807-263-4
; Sequence 4, Application US/08807263C
; Patent No. 5985627
; GENERAL INFORMATION:
; APPLICANT: Mortensen, Uffe
; APPLICANT: Olesen, Kjeld
; APPLICANT: Stennicke, Henning
; APPLICANT: Sorensen, Steen B.
; APPLICANT: Breddam, Klaus
; TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE
; FILE REFERENCE: B648.71us01-no4
; CURRENT APPLICATION NUMBER: US/08/807,263C
; CURRENT FILING DATE: 1997-02-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae

US-08-807-263-4
Alignment Scores:
Pred. No.: 21.7 Length: 421
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 2 Gaps: 0

US-09-712-338-1 (1-1668) x US-08-807-263-4 (1-421)
QY 298 TTGTGTTGATGCTGGCCCTGGA 321
DB 48 LeuTrpLeuAsnGlyGlyProGly 55

RESULT 13
US-08-828-488-8
; Sequence 8, Application US/08828488
; Patent No. 5925521
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN SERINE
; TITLE OF INVENTION: CARBOXYPEPTIDASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,488
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0241 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 190283
US-08-828-488-8

Alignment Scores:
Pred. No.: 21.5 Length: 480
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 2 Gaps: 0
```

US-09-712-338-1 (1-1668) x US-09-828-488-8 (1-480)

Qy 298 TTGTGGTTGAATGGTGGCCCTGGA 321
Db 80 LeuTrpLeuAsnGlyGlyProGly 87

RESULT 14

US-09-299-689A-8

; Sequence 8, Application US/09299689A

; Patent No. 6379913

GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Hawkins, Phillip R.

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Lal, Preeti

; APPLICANT: Goli, Surya K.

; TITLE OF INVENTION: NOVEL HUMAN SERINE

; TITLE OF INVENTION: CARBOXYPEPTIDASE

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/299,689A

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/828,488

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0241 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 480 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GenBank

; CLONE: 190283

US-09-299-689A-8

Alignment Scores:
Pred. No.: 21.5 Length: 480
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 4 Gaps: 0

US-09-712-338-1 (1-1668) x US-09-299-689A-8 (1-480)

Qy 298 TTGTGGTTGAATGGTGGCCCTGGA 321
Db 80 LeuTrpLeuAsnGlyGlyProGly 87

RESULT 15

US-09-702-705-336

; Sequence 336, Application US/09702705

; Patent No. 6504010

GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-705-336

Alignment Scores:
Pred. No.: 21.5 Length: 480
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.44% Indels: 0
DB: 4 Gaps: 0

US-09-712-338-1 (1-1668) x US-09-702-705-336 (1-480)

Qy 298 TTGTGGTTGAATGGTGGCCCTGGA 321
Db 80 LeuTrpLeuAsnGlyGlyProGly 87

Search completed: November 21, 2003, 17:56:40
Job time : 25.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:40:04 ; Search time 63.5 Seconds
(without alignments)
9590.842 Million cell updates/sec

Title: US-09-712-338-1

Perfect score: 3027

Sequence: 1 atcgctggctacgaattctt.....ccagtggtgtgtagcatag 1668

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Zgapop 6.0 , Zgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 666188 seqs, 18259486 residues

Total number of hits satisfying chosen parameters: 1332376

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODELS=frame+ n2p.model -DEV=xlp

-Q=/cgn2_1/USPTO.spool.p/US09712338/runat 17112003 170150 10599/app_query.fasta_1.1863

-DB=Published Applications AA -OPMT=fastn -SUFFIX=rapb -MINMATCH=0.1

-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62

-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100

-THR_MIN=0 -ALIGN=15 -MODES=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=2000000000 -USER=US09712338 @CGN 1 1 15 @runat 17112003 170150 10599

-NCFU=6 -ICPU=3 -NO MAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100

-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	517.5	17.1	421	9	US-09-420-785A-4	Sequence 4, Appli
2	517.5	17.1	421	9	US-09-901-252-15	Sequence 15, Appl
3	361.5	11.9	476	14	US-10-084-018-3	Sequence 3, Appli
4	361.5	11.9	476	10	US-09-909-320-164	Sequence 164, App
5	360.5	11.9	476	10	US-09-909-088B-164	Sequence 164, App
6	360.5	11.9	476	10	US-09-905-291A-164	Sequence 164, App
7	360.5	11.9	476	10	US-09-902-853-164	Sequence 164, App
8	360.5	11.9	476	10	US-09-907-824-164	Sequence 164, App
9	360.5	11.9	476	10	US-09-907-841-164	Sequence 164, App
10	360.5	11.9	476	11	US-09-904-011-164	Sequence 164, App
11	360.5	11.9	476	11	US-09-906-742-164	Sequence 164, App
12	360.5	11.9	476	11	US-09-906-838-164	Sequence 164, App
13	360.5	11.9	476	11	US-09-907-613-164	Sequence 164, App
14	360.5	11.9	476	11	US-09-907-942-164	Sequence 164, App
15	360.5	11.9	476	11	US-09-796-753-40	Sequence 40, Appl
16	360.5	11.9	476	11	US-09-904-859-164	Sequence 164, App
17	360.5	11.9	476	11	US-09-909-204-164	Sequence 164, App
18	360.5	11.9	476	11	US-09-904-820-164	Sequence 164, App
19	360.5	11.9	476	11	US-09-904-786-164	Sequence 164, App
20	360.5	11.9	476	11	US-09-906-646-164	Sequence 164, App
21	360.5	11.9	476	11	US-09-906-700-164	Sequence 164, App
22	360.5	11.9	476	11	US-09-903-786-164	Sequence 164, App
23	360.5	11.9	476	11	US-09-902-903-164	Sequence 164, App
24	360.5	11.9	476	11	US-09-903-749A-164	Sequence 164, App
25	360.5	11.9	476	11	US-09-904-119-164	Sequence 164, App
26	360.5	11.9	476	11	US-09-904-956-164	Sequence 164, App
27	360.5	11.9	476	11	US-09-902-736-164	Sequence 164, App
28	360.5	11.9	476	11	US-09-907-794-164	Sequence 164, App
29	360.5	11.9	476	11	US-09-903-943-164	Sequence 164, App
30	360.5	11.9	476	11	US-09-904-462-164	Sequence 164, App
31	360.5	11.9	476	11	US-09-907-925-164	Sequence 164, App
32	360.5	11.9	476	11	US-09-902-692-164	Sequence 164, App
33	360.5	11.9	476	11	US-09-903-520-164	Sequence 164, App
34	360.5	11.9	476	11	US-09-905-056-164	Sequence 164, App
35	360.5	11.9	476	11	US-09-909-064-164	Sequence 164, App
36	360.5	11.9	476	11	US-09-904-553-164	Sequence 164, App
37	360.5	11.9	476	11	US-09-905-381-164	Sequence 164, App
38	360.5	11.9	476	11	US-09-905-088-164	Sequence 164, App
39	360.5	11.9	476	11	US-09-907-575-164	Sequence 164, App
40	360.5	11.9	476	11	US-09-905-075-164	Sequence 164, App
41	360.5	11.9	476	11	US-09-902-759-164	Sequence 164, App
42	360.5	11.9	476	11	US-09-902-634-164	Sequence 164, App
43	360.5	11.9	476	11	US-09-907-979-164	Sequence 164, App
44	360.5	11.9	476	11	US-09-902-615-164	Sequence 164, App
45	360.5	11.9	476	11	US-09-902-615-164	Sequence 164, App

ALIGNMENTS

RESULT 1

US-09-420-785A-4

; Sequence 4, Application US/09420785A

; Patent No. US20010010923A1

; GENERAL INFORMATION:

; APPLICANT: MORTENSEN, UFFE

; APPLICANT: OLESEN, KJELD

; APPLICANT: STENNICK, HENNING

; APPLICANT: SORESEN, STEEN B.

; APPLICANT: BREDDAM, KLAUS

; TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE

; FILE REFERENCE: 089187/0109

; CURRENT APPLICATION NUMBER: US/09/420,785A

; CURRENT FILING DATE: 1999-10-19

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 421

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-09-420-785A-4

Alignment Scores: 2.22e-34 Length: 421

Pred. No.: 421

QY 268 AGACATAACCCAGAACTGCACATATACATTTGTTGAATGGTGGCCCTGGAAGCAT 327
 Db |||||:|||||
 QY 38 ArgAsnAspProAlaAlaAspProValIleLeuThrLeuAsnGlyGlyProGlyCysSer 57
 Db |||||:|||||
 QY 328 TCTTTGATCGGTCTCTTCAAGAGTGGCCCTGCCATGTCAATTCGACTTTTGATGAC 387
 Db |||||:|||||
 QY 58 SerLeuThrGlyLeuPheGluLeuGlyProSerSerIleGlyProAspLeuLysPro 77
 Db |||||:|||||
 QY 388 TACATCAACCTCACTCGTGGAAACGAGGTCTCAATTTACTATTCTCTCCAGCATTTG 447
 Db |||||:|||||
 QY 78 IleGlyAsnProTyrSerTrpAsnSerAsnAlaThrValIlePheLeuAspGlnProVal 97
 Db |||||:|||||
 QY 448 GGAGTCGGCTTTTCATATAGTGAATACAGGTGATGGGTCCATTAAACCTGTAACTGGGGTC 507
 Db |||||:|||||
 QY 98 AsnValGlyPheSerTyrSer----- 104
 QY 508 GTCGAAAATTCGAGCTTTGCGAGGAGTTGAGGGCCGTTACCCAACATTGATGCCACTGTG 567
 Db |||||:|||||
 QY 105 -----GlySerSerGly----- 108
 QY 568 ATCGATACCTACCAATCTTTCGCGCAGAGCCGCTTGGGAGATCTCTCAAGGATTCCTTAGT 627
 Db |||||:|||||
 QY 109 ValSerAsnThrValAlaAlaGlyAspValTyrAsnPheLeuGluLeuPhePheAsp 128
 Db |||||:|||||
 QY 628 GGACTACTAGCTGGACTTAGGGTGCAGTCTAAGGACTTCACTCTATGTCAGCGAGGC 687
 Db |||||:|||||
 QY 129 GlnPheProGluTyrValAsnLysGlyGln-----AspPheHisIleAlaGlyGluSer 146
 Db |||||:|||||
 QY 688 TATGGAGCGCACTATGCTCGCTGCAATCTTCATCATTTTACGAGCAGAAATGAGAGAATT 747
 Db |||||:|||||
 QY 147 TyrAlaGlyHisTyrIleProValPheAlaSerGluLeuSerHisLysAspArg--- 165
 QY 748 GCCAACGGTAGTGTAAATGGTGTTCAGCTTAATTTCAACTCTCTGGGAATTT--- 801
 Db |||||:|||||
 QY 166 -----AspPheAsnLeuThrSerValLeuIleGly 175
 QY 802 AACGGCATCATCGACGAGCGATCCAGGCCCTTACTTACCTCGAATTCGCT----- 852
 Db |||||:|||||
 QY 176 AsnGlyLeuThrAspProLeuThrGlnTyrAsnTyrTyrGluProMetAlaCysGlyGlu 195
 Db |||||:|||||
 QY 853 -----GTGACAATACCTAGCTATCAGCTGTACAGCTGTCAACGAGACCGTCTAC 897
 Db |||||:|||||
 QY 196 GlyGlyGluProSerValLeuProSerGluGluCysSerAlaMetGluAspSerLeu--- 214
 Db |||||:|||||
 QY 898 AACTACATGAAGTTTGCCACCAATGCCAAATGCTGCCAGATTTGATTTCACCTGC 957
 Db |||||:|||||
 QY 215 -----GluArgCysLeuGlyLeuIleGluSerCys 224
 QY 958 AAACACACAAACCGCAGCATTTAGCTGACTAGCCCTCTGCGCCGGAAGCCACCAACATG 1017
 Db |||||:|||||
 QY 225 TyrAspSerGln-----SerValTrpSerCysValProAlaThrIleTyr 239
 QY 1018 TGCAAGGACAAATGTTGAGGGCCCATACTACGCTTTGCTGGTGTGGTGTGATGATTT 1077
 Db |||||:|||||
 QY 240 CysAsnAsnAlaGlnLeuAlaProTyrGlnArg---ThrGlyArgAsnValTyrAspIle 258
 Db |||||:|||||
 QY 1078 CGCATCCATATGATGAC-----CCGACTCCGCAAGTTATTACACAAA 1122
 Db |||||:|||||
 QY 259 ArgLysAspCysGluGlyGlyAsnLeuCysTyrProThrLeuGlnAsp---IleAspAsp 277
 Db |||||:|||||
 QY 1123 TTTCTGGCAAGAGCTCTCTCATGGACGTATGGCGTCAACATCACTACACCCAGTCC 1182
 Db |||||:|||||
 QY 278 TyrLeuAsnGlnAspTyrValIysGluAlaValGlyAlaGluValAspHisTyrGluSer 297
 Db |||||:|||||
 QY 1183 ---AATAATGAGCTCTACTACGCTTCCAGCAAAACAGGCGACTTGTCTGGCCCACTTC 1239
 Db |||||:|||||
 QY 298 CysAsnPheAspIleAsnArgAsnPheLeuPheAlaGlyAspTrpMetLysPro---Tyr 316
 Db |||||:|||||
 QY 1240 ATCGAAGACCTCGAGGAGATCTTGTCTCTCCCGTGGTGTCTCTCATCTATCGCGAC 1299
 Db |||||:|||||
 QY 317 HisThrAlaValThrAspLeuLeuAsnGlnAspLeuProIleLeuValTyrAlaGlyAsp 336

QY 1300 GCCGATTACATCTGCAACTGTTGCGCGGTGAGCGGTTTCCCTGCTGCGAACTACTCC 1359
 Db |||||:|||||
 QY 337 LysAspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrAspValLeuProTrpLys 356
 Db |||||:|||||
 QY 1360 CAAGCGGCCAGTTTCCGAAGC-----GCAGGGTACAGCCCTGAAAGTCAACGCGC 1410
 Db |||||:|||||
 QY 357 TyrAspGluGluPheAlaSerGlnLysValArgAsnTrpThrAlaSerIleThrAspGlu 376
 Db |||||:|||||
 QY 1411 GTCGAGTATGGGAAACTCGAGATATGGTAATTTCTCTTCACTCGCGGTCTATGAGCA 1470
 Db |||||:|||||
 QY 377 ValAla---GlyGluValLysSerTyrLysHisPheThrTyrLeuArgValPheAsnGly 395
 Db |||||:|||||
 QY 1471 GGCCATGAGTCACCATCTACTACAGCCCATCCCTCCCTGCAATTTTAAACCGGACTATC 1530
 Db |||||:|||||
 QY 396 GlyHisMetValProPheAspValProGluAsnAlaLeuSerMetValAsnGluTrpIle 415
 Db |||||:|||||
 QY 1531 TTCGGT 1536
 Db |||||:|||||
 QY 416 HisGly 417
 Db |||||:|||||
 RESULT 3
 US-10-084-018-3
 ; Sequence 3, Application US/10084018
 ; Publication No. US20020160499A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; Hawkins, Phillip R.
 ; Hillman, Jennifer L.
 ; Lal, Preeti
 ; Goli, Survi K.
 ; TITLE OF INVENTION: NOVEL HUMAN SERINE
 ; CARBOXYPEPTIDASE
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/084,018
 ; FILING DATE: 25-Feb-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/299,689A
 ; FILING DATE: <Unknown>
 ; APPLICATION NUMBER: 08/828,488
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0241 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 476 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: MMLR3DT01
 ; CLONE: 566993
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 US-10-084-018-3

Alignment Scores:

Pred. No.:	2,41e-21	Length:	476
Score:	361.50	Matches:	141
Best local Similarity:	29.03%	Conservative:	69
Query Match:	11.94%	Mismatches:	201
DB:	14	Indels:	128
		Gaps:	23
US-09-712-338-1 (1-1668) x US-10-084-018-3 (1-476)			
QY	50	GGGCGCTCCAGGAAGTACACGGCGTGGTGGTAGAGACAGCTACCCAGAACCCCA	109
DB	22	GlyLeuPheHisSerLeuTyrArgSerValSerMet-----Pro	34
QY	110	CGGGGTCAAGACTCTTACAAACCGAAACAATGTACCA-----TCCGGTACAAGG	160
DB	35	ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle	54
QY	161	AACCCGGGCGACAGGGCGTCTCGAGACTACCCCG-----GGTCCAAATCTTAC	210
DB	55	GlnLysGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr	74
QY	211	TCGGATAT-----GTCCACACTCTCCGAGTCCATACCTTCTCTGGTTCTTCGAA	264
DB	75	AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPheThrPhePro	94
QY	265	GCAGACATAACCCAGAACTGCACCTATCATCAATTGTGGTTGAATGGTGGCCCTGGAAGC	324
DB	95	AlaGlnIleGlnProGluAspAlaProValValLeuTripleGlnGlyProGlyGly	114
QY	325	GATTCCTTCGCTCTTCGAGAGAGTTGGGCGCTTGGCATGTCAATTCGACTTTTGAT	384
DB	115	SerSerMet**GlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr	134
QY	385	GACTACATCAACCTCACTCGTGGACAGAGGTCTCCAAATTACTATTCTGTGCCAGCCA	444
DB	135	LeuArgAspArgAspPheProThrThrThr***SerMetLeuIleAspAsnPro	154
QY	445	TTGGGAGTCCGGCTTTTCATATAGTGTATACGGTTGATGGGTCATTAAACCTGTAACTGGG	504
DB	155	ValGlyThrGlyPheSerPheThrAspAspThrHisGly-----	167
QY	505	GTCTCGAAATTCGAGCTTTGCGAGGTTTCAGGGGCGGTACCCACCACTGTGCACT	564
DB	168	-----Tyr	168
QY	565	CTGATCGATACCTATCTTGGCGAGAGCGCGCTTGGGAGATCTCGCAAGGATTCCTT	624
DB	169	AlaValAsnGluAspAspValAlaArgAspLeuTyrSerAlaIleIleGlnPhePhe---	187
QY	625	AGTGGACTACCTAGCTTGGACTCTAGGGTGCAGTCTAAGGACTTCAGTCTATGGAACGAG	684
DB	188	-----GlnIlePheProGluTyrLysAsnAsnAspPheTyrValThrGlyGlu	203
QY	685	AGTATGGAGGACATATGCTGCTGCAATCTTCAATCATTTTACGAGCAGATGAGAGA	744
DB	204	SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn-----	221
QY	745	ATTGCCAACGGTAGTGTAAATGGTGTTCAGCTTAAATTCCTCTCTGGGAATTTATAAC	804
DB	222	-----ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp	237
QY	805	GGCATATCGACGAGGCGATCCAGCGCCCTTACTACCTGAAATTCGCTGTGAAACAATACC	864
DB	238	GlyTyrSerAspProGluSerIleIleGlyTyrAlaGluPhe-----Leu	253
QY	865	TACGGTATCAGGCTGTCAACGAGACCGCTCTACAACTACATGAAGTTTGGCAACCAATG	924
DB	254	TyrGlnIleGlyLeuLeuAspGluLysGlnLysLysTyr-----PheGlnLysGln---	270
QY	925	CCAAATGTTTCCAGGAGTTTCAATTTCCACTTGCACCAACAGACAAAAC-----CGCACC	975
DB	271	-----CysHisGluCysIleGluHisIleArgLysGlnAsnTyrPheGluAlaPhe	287

RESULT 4
 US-09-309-320-164
 ; Sequence 164, Application US/09909320
 ; Patent No. US20020132240A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, A.
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth, J.
 ; APPLICANT: Kijavina, Ivar J.
 ; APPLICANT: Mather, Jennie P.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: 10466-14
 ; CURRENT APPLICATION NUMBER: US/09/909,320
 ; PRIOR FILING DATE: 2002-01-04
 ; PRIOR APPLICATION NUMBER: PCT/US00/04414
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: US 60/143,048
 ; PRIOR FILING DATE: 1999-07-07
 ; PRIOR APPLICATION NUMBER: US 60/145,698
 ; PRIOR FILING DATE: 1999-07-26
 ; PRIOR APPLICATION NUMBER: US 60/146,222
 ; PRIOR FILING DATE: 1999-07-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/20594
 ; PRIOR FILING DATE: 1999-09-08
 ; PRIOR APPLICATION NUMBER: PCT/US99/20944
 ; PRIOR FILING DATE: 1999-09-13
 ; PRIOR APPLICATION NUMBER: PCT/US99/21090
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/21547
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/23089
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: PCT/US99/28214
 ; PRIOR FILING DATE: 1999-11-29
 ; PRIOR APPLICATION NUMBER: PCT/US99/28313
 ; PRIOR FILING DATE: 1999-11-30
 ; PRIOR APPLICATION NUMBER: PCT/US99/28564
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/28565
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/30095
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: PCT/US99/30911
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US99/30999
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US00/00219
 ; PRIOR FILING DATE: 2000-01-05
 ; NUMBER OF SEQ ID NOS: 423
 ; SEQ ID NO 164
 ; LENGTH: 476
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-909-320-164

Alignment Scores:
 Pred. No.: 2,92e-21 Length: 476
 Score: 360.50 Matches: 141
 Percent Similarity: 39.22% Conservative: 70
 Best Local Similarity: 26.21% Mismatches: 200
 Query Match: 11.91% Indels: 128
 DB: 10 Gaps: 23

US-09-712-338-1 (1-1668) x US-09-909-320-164 (1-476)

QY	50	GGGCGCTCCAGGAGTACACCGGCTCGGTGGTAGAGACAGTACCCAGAACCCCA	109
Db	22	GlyLeuPheArgSerLeuTyrArgSerValSerMet	34
QY	110	CGGGGTCAAGACTTTACACCGCAACAAATGTCCACCA	160
Db	35	ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle	54
QY	161	AACCCGGCGCAGGGCGTCTCGGAGCTACCCCG	210
Db	55	GlnLysGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr	74
QY	211	TCGTGATAT	264
Db	75	AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPhePheTrpPhePro	94

QY	265	GCCAGACATACCCAGAACTGCACCTATCATCTGTGGTGAATGTGGCCCTGGAACG	324
Db	95	AlaGlnIleGlnProGluAspAlaProValValLeuTrpLeuGlnGlyGlyProGlyGly	114
QY	325	GATCTTTGATCGGTCTCTCGAAGAGTTGGCCCTTGCACATGTCATTCACATTTTCAT	384
Db	115	SerSerMetPheGlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr	134
QY	385	GACTACATCAACCCCTCACTCGTGAACGAGGTCTCCCAATTTACTATTCTCTCCAGCCA	444
Db	135	LeuArgAspArgAspPheProTyrThrThrLeuSerMetLeuTyrIleAspAsnPro	154
QY	445	TTGGAGTCGGCTTTTCATATAGTGATACGGTTCATGGTCCATTAACCCCTGTAACCTGGG	504
Db	155	ValGlyThrGlyPheSerPheThrAspAspThrHisGly	167
QY	505	GTCTCGAAAAATTCGAGCTTTCGAGGAGTTTCAGGGCCGTTACCCCAACCATGTATGCCACT	564
Db	168		Tyr 168
QY	565	CTGATGATACTACCAATCTTCCGAGAGCGCTTGGGAGATCTCTCAAGGATTCTCT	624
Db	169	AlaValAsnGluAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe	187
QY	625	AGTGGACTACTAGCTTGGACTTAGGTGCTAGTCTAAGGACTTCAGTCTATGACGAGGAG	684
Db	188		GlnIlePheProGluTyrLysAsnAspPheTyrValThrGlyGlu 203
QY	685	AGCTATGGAGGCACTATGCTCTGCAATCTTCAATCATTTTACGAGCAGAGATGAGAGA	744
Db	204	SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn	221
QY	745	ATTGCCAAGGTAGTCTTAATGTTGTTCAGCTTAATTTCACTCTCTGGAATATTATAC	804
Db	222		ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp 237
QY	805	GGCATCATCGAGGCGATCCAGGCCCTTACTACCTGAATTCGCTGTGAACAATATG	924
Db	238	GlyTyrSerAspProGluSerIleIleGlyGlyTyrAlaGluPhe	924
QY	865	TACGGTATCAAGGCTGTCAACGAGACCGTCTACACATACATGAGTTGCCAACCAATG	924
Db	254	TyrGlnIleGlyLeuLeuAspGlnLysGlnLysLysTyr	270
QY	925	CCAAATGGTTGCCAGGATTGATTTCCACCTGCACACAGACAAAC	975
Db	271		CysHisGluCysIleGluHisIleArgLysGlnAsnTrpPheGluAlaPhe 287
QY	976	GCATTAGCTGACTACGCCCTCTGCGCGGAGCCCAAC	1029
Db	288	GluIleLeuAspLysLeuLeuAspGlyAspLeuThrSerAspProSerTyrPheGlnAsn	307
QY	1030	GTTCAGGGG	1083
Db	308	ValThrGlyCysSerAsnTyrTyrAsnPheLeu	318
QY	1084	CCATATGATGACCCGACTCCGCCAAGTATTACAAATAATTTCTGCAAGGACTCTGTC	1143
Db	319	ArgCysThrGluProGluAspGlnLeuTyrTyrValLysPheLeuSerLeuProGluVal	338
QY	1144	ATGGAGCTATCGGCTCAACATCACTACACCCAGTCCAAATGACGTCTAC	1197
Db	339	ArgGlnAlaIleHisValGly	357
QY	1198	TACGCTTTCCAGCAACAGGCGACTTTCTGTCGGCCCACTTCATCGAGACCTCGAGGAG	1257
Db	358	TyrLeuArgGluAspThrValGlnSerValLysProTyrLeuThrGluIleMetAsnAsn	377
QY	1258	ATCCTTGCTCTCCCGTCTCGTGTCTCCCTCATCTATGCGAGCGCGGATTACATC	1311
Db	378		TyrLysValIleIleTyrAsnGlyGlnLeuAspIleIleValAla 392
QY	1312		TGCAACTGTTCCGGCGGTACGGCCGTTTCC 1341

```

; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-088B-164

```

RESULT 5

```

US-09-909-088B-164
; Sequence 164, Application US/09909088B
; Patent No. US20020146702A1

```

GENERAL INFORMATION:

```

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fillvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,088B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214

```

```

; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-088B-164

Alignment Scores:
Pred. No.: 2,92e-21 Length: 476
Score: 360.50 Matches: 141
Percent Similarity: 39.22% Conservative: 70
Best Local Similarity: 26.21% Mismatches: 200
Query Match: 11.91% Indels: 128
DB: 10 Gaps: 23

US-09-712-338-1 (1-1668) x US-09-909-088B-164 (1-476)
QY 50 GGGCCCTCCAGAGAGTACACCGGCGTCGTCGGTAGAGACAGCTACCCAGAACCCCA 109
Db 22 GlyLeuPheArgSerLeuTyrArgSerValSerMet-----Pro 34
QY 110 CCGGGGTCAAGACTCTTACACCGCAACAACTGTACCA-----TCCGTACAAGG 160
Db 35 ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysile 54
QY 161 AACCCGGGCGAGAGGCGTCTCGGAGACTACCCCG-----GGTGTCAATCTCTAC 210
Db 55 GlnLysGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr 74
QY 211 TCTGGATAT-----GTGACACCTCTCCGAGTCCCATACCTCTCTGGTCTTCGAA 264
Db 75 AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPheThrPhePro 94
QY 265 GCCAGACATACCCAGAACTGCACCTATCACATTTGTGGTTGAATGGTGGCCCTGGAAGC 324
Db 95 AlaGlnIleGlnProGluAspAlaProValValLeuTyrPleuGlnGlyGlyProGlyGly 114
QY 325 GATTCTTTGATCGGTCTCTTCGAAGAGTTGGGCGCTTGGCCATGTCAATTCGACTTTGAT 384
Db 115 SerSerMetPheGlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr 134
QY 385 GACTATCATCAACCCCTCACTCGTGAACGAGGTCTCCAAATTTACTATTCTCTGTCACGCCA 444
Db 135 LeuArgAspArgAspPheProTyrThrThrLeuSerMetLeuTyrIleAspAsnPro 154
QY 445 TTGGGAGTCGGCTTTTCATATAGTATACGCTTGTGATGGTTCATTAACCCCTGTAACTGGG 504
Db 155 ValGlyThrGlyPheSerPheThrAspAspThrHisGly----- 167
QY 505 GTCGTGAAAATTCGAGCTTTGCAGGAGTTTCAGGGCGGTACCCACCATGTATGCCACT 564
Db 168 -----Tyr 168
QY 565 CTGATCGATACCAATCTTCGCGAGAGCGCCCTTGGGAGATCTCTGCAAGGATCTCTT 624
Db 169 AlaValAsnGluAspAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe--- 187
QY 625 AGTGGACTACCTAGCTTGGACTCTTAGGGTGCAGTCTAAGGACTTCAAGGACTTCAATGACGGAG 684

```

Db 188 -----GlnIlePheProGluTyrLysAsnAsnAspPheTyrValThrGlyGlu 203
QY 685 AGCTATGAGGCACTATGGCTCTCAATCTTCAATCTTACGAGCAGATGAGAGA 744
Db 204 SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn 221
QY 745 ATTGCGCAACGGTAGTGTAAATGGTTCAGCTTAATTTCAACTCTCTGGGAATATTAAAC 804
Db 222 -----ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp 237
QY 805 GGCATATCGACGAGGCGATCCAGCCCTTACTACCTCGATTCGCTGTGAACAATACC 864
Db 238 GlyTyrSerAspProGluSerIleIleGlyGlyTyrAlaGluPhe-----Leu 253
QY 865 TACGGTATCAAGGCTGTCAACAGCAGCCGTCAACACTACAGTTCGCAACCAATG 924
Db 254 TyrGlnIleGlyLeuLeuAspGluLysGlnIlystYr-----PheGlnLysGln 270
QY 925 CCAATGGTTCGCAAGGATTGATTTCCACCTGCAACAGACAAC-----CGCAC 975
Db 271 -----CysHisGluCysIleGluHisIleArgLysGlnAsnTyrPheGluAlaPhe 287
QY 976 GCATTAGCTGACTACGCGCTCTGCGCCGAGCACCACCAAC-----ATGTGAGGGAACAAT 1029
Db 288 GluIleLeuAspLysLeuLeuAspGlyAspLeuThrSerAspProSerTyrPheGlnAsn 307
QY 1030 GTTAGGG-----CCATACTACGCTCTGCTGCTGTGTGTATGATATTCGGCAT 1083
Db 308 ValThrGlyCysSerAsnTyrTyrAsnPheLeu----- 318
QY 1084 CCATATGATACCGGACTCGCCCAAGTTATTACACAAATTTCTGCAAAAGGACTCTGTC 1143
Db 319 ArgCysThrGluProGluAspGlnLeuTyrTyrValLysPheLeuSerLeuProGluVal 338
QY 1144 ATGACGCTATCGCGTCAACATCACTACACCGACTCCCAATATGACGCTAC----- 1197
Db 339 ArgGlnAlaIleHisValGly-----AsnGlnThrPheAsnAspGlyThrIleValGluLys 357
QY 1198 TAGCTTTTCCAGCAACAGCGCACTTGTCTGCGCCCACTTCATCGAAGACCTCGAGGAG 1257
Db 358 TyrLeuArgGluAspThrValGlnSerValLysProTyrLeuThrGluIleMetAsnAsn 377
QY 1258 ATCCTTGCTCTCCGCTGCTGCTCTCCCTCATCTATGGCAGCGCGATTACATC----- 1311
Db 378 -----TyrLysValLeuIleTyrAsnGlyGlnLeuAspIleIleValAla 392
QY 1312 -----TGCAACTGTTCCGGCGGTTCAGCGCGTTCAGTGGGAAACTCCCGAG 1341
Db 393 AlaAlaLeuThrGluArgSerLeuMetGlyMetAspTyrLysGlySerGln----- 409
QY 1342 CTGCTGCGAACTACTCCCAAGCGCCCGAG-----TTCCGAAGC----- 1380
Db 410 -----GluTyrLysLysAlaGluLysLysValTyrLysIlePheLysSerAspSer 426
QY 1381 -----GAGGGTACAGCCCTCGAAGTCAAGCGGCTCGAGTATGGGAAACTCCCGAG 1434
Db 427 GluValAlaGlyTyr-----IleArgGln 434
QY 1435 TATGGTAATTTCTCCTTCACTCGCTCTATGAGGCGGCGCATGAGTCCATCTACACGAG 1494
Db 435 AlaGlyAspPheHisGlnValIleIleArgGlyGlyGlyHisIleLeuProTyrAspGln 454
QY 1495 CCCATCGCTCCCTGCAATGTGTTAACCGGACTATCTTC-----GGTTGGGAT 1542
Db 455 ProLeuArgAlaPheAspMetIleAsnArgPheIleTyrGlyLysGlyTyrAsp 472

RESULT 6

US-09-905-291A-164
; Sequence 164, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavan, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-164
Alignment Scores: 2,92e-21 Length: 476
Pred. No.: 360.50 Matches: 141
Score:

Percent Similarity: 39.22% Conservative: 70
 Best Local Similarity: 26.21% Mismatches: 200
 Query Match: 11.91% Indels: 128
 DB: 1.0 Gaps: 23

US-09-712-338-1 (1-1668) x US-09-905-291A-164 (1-476)

50 GGGCCCTCCAGGAGTACACCGCGTCCGTCGGTAGAAGACAGCTACCCCAAGACCCCA 109
 22 GlyLeuPheArgSerLeuTyArgSerValSerMet-----Pro 34
 110 CGGGGTCAAGACTCTTACAACCGCAAAATGTCCACCA-----TCGGTACAAGG 160
 35 ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyRileGluAlaGlyLysile 54
 161 AACCCGGGACAGGGCGTCTCGAGACTACCCCG-----GGTGCAAAATCCCTAC 210
 55 GlnTyGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyR 74
 211 TCTGATAT-----GTGCACACCTCTCCGAGTCCCATACCTCTCTCTGTTCTTCGAA 264
 75 AlaGlyPheLeuThrValAsnLysThrTyRAsnSerAsnLeuPhePheTrpPhePro 94
 265 GCAGACATAACCCAGAACTGCATATCATCTGTGTGTTGAATGGTGGCCCTCGAAGC 324
 95 AlaGlnIleGlnProGluAspAlaProValValLeuTrpLeuGlnGlyProGlyGly 114
 325 GATCTTTGATCGTCTCTCGAAGAGTGGCCCTGCATCTCAATTCGACTTTTGAT 384
 115 SerSerMetPheGlyLeuPheValGluHisGlyProTyRValValThrSerAsnMetThr 134
 385 GACTACATCAACCTCACTCGTGAAGAGGTCTCCAAATTTACTATTCTCTGCCCA 444
 135 LeuArgAspArgAspPheProTrpThrThrLeuSerMetLeuTyRileAspAsnPro 154
 445 TTGGGAGTCGGCTTTTCATATAGTATGATACGTTGATGGTCCATTAACCTGTAACTGG 504
 155 ValGlyThrGlyPheSerPheThrAspThrHisGly----- 167
 505 GTCGTGAATAATCGAGCTTTGCAGAGTTTCAGGGCCGGTACCCCAACCATTTGATGCCACT 564
 168 -----Tyr 168
 565 CTGATCGATACTACCAATCTTGGCAGAGCGCTTGGAGATCTCTGCAAGATTCCTT 624
 169 AlaValAsnGluAspValAlaArgAspLeuTyRSerAlaLeuIleGlnPhePhe--- 187
 625 AGTGAATCACTAGCTTGGACTCTAGGGTCAGTCTAAGACTTCAGTCTATGGACGAG 684
 188 -----GlnIlePheProGluTyRAsnAsnAspPheTyRValThrGlyGlu 203
 685 AGCTATGAGGCACTATGCTCTGCTGATCTTCAATCATTTTACGAGCAGATGAGAGA 744
 204 SerTyRAlaGlyLysTyRValProAlaIleAlaHisLeuIleHisSerLeuAsn--- 221
 745 ATTGCCAACCGTAGTGTAAATGTTGCTTCACTTAATTCACCTCTCTGGGAATTTAATC 804
 222 -----ProValArgGluValLysileAsnLeuAsnGlyIleAlaIleGlyAsp 237
 805 GGATATATCGACGAGGATCCAGGCCCTTACTACCTGATTCGCTGTTGAAACAATACC 864
 238 GlyTyRSerAspProGluSerIleIleGlyTyRAlaGluPhe-----Leu 253
 865 TACGGTATCAAGGCTGTCAACGAGACCGTCTACAACTACATGAAATTTGCCAACCAATG 924
 254 TyrGlnIleGlyLeuLeuAspGluLysLysTyR-----PheGlnLysGln--- 270
 925 CCAAAATGTTCCAGGATTTGATTTCCACCTGCAACAGACAAAC-----CGCACC 975
 271 -----CysHisGluCysIleGluHisIleArgLysGlnAsnTrpPheGluAlaPhe 287
 976 GCATTAGTCTAGTACGCTCTCGCCGAGGACCAAC-----ATGTGACGAGGACAAAT 1029

288 GluIleLeuAspLysLeuLeuAspGlyAspLeuThrSerAspProSerTyR-PheGlnAsn 307
 1030 GTTGAGGG-----CCATACTACGCTTCTCTGCTGGTGTGTATGATATTCGGCAT 1083
 308 ValThrGlyCysSerAsnTyRAsnPheLeu----- 318
 1084 CCATATGATGACCGGACTCCGCCAAGTTATTACAAATAATTTCTGGCAAGGACTCTCTC 1143
 319 ArgCysThrGluProGluAspGlnLeuTyRValLysPheLeuSerLeuProGluVal 338
 1144 ATGACAGCTATCGGCGTCAACATCACTACACCCAGTCCCAATAATGAGCTCTAC----- 1197
 339 ArgGlnAlaIleHisValGly---AsnGlnThrPheAsnAspGlyThrIleValGluLys 357
 1198 TACGCTTTCCAGCAACAGGCGACTTTGCTGGCCCACTTCATCGAAGACCTCGAGGAG 1257
 358 TyrLeuArgGluAspThrValGlnSerValLysProTrpLeuThrGluLeuMetAsnAsn 377
 1258 ATCCTTGCTCTCCCGTGTCTCCCTCATCTATGCGACGCGCATTTACATC----- 1311
 378 -----TyrLysValLeuIleTyRAsnGlyGlnLeuAspIleIleValAla 392
 1312 -----TGCAACTGTTCCGCGGTGAGCGGCTTCCGAAAGC----- 1341
 393 AlaAlaLeuThrGluArgSerLeuMetGlyMetAspTrpLysGlySerGln----- 409
 1342 CTCGCTGCGAACTATCCCAAGCCGCCAG-----TTCGGAAGC----- 1380
 410 -----GluTyRlysAlaGluLysLysValTrpLysIlePheLysSerAspSer 426
 1381 -----GCAGGTCACAGCCCTCGAAAGTCAACGGCGTTCGATGGGAAACTCGCGAG 1434
 427 GluValAlaGlyTyR----- 434
 1435 TATGTAATTTCT 1494
 435 AlaGlyAspPheHisGlnValIleIleArgGlyGlyHisIleLeuProTyRAspGln 454
 1495 CCCATCGCT 1542
 455 ProLeuArgAlaPheAspMetIleAsnArgPheIleTyRglyLysGlyTrpAsp 472

RESULT 7

US-09-902-853-164
 ; Sequence 164, Application US/09902853
 ; Publication No. US20020192659A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, A.
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth, J.
 ; APPLICANT: Kijavain, Ivar J.
 ; APPLICANT: Mather, Jennie P.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William, I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic


```

; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-712-338-1 (1-1668) x US-09-712-338-1 (1-476)

Alignment Scores:
Pred. No.: 2,92e-21 Length: 476
Score: 360.50 Matches: 141
Percent Similarity: 39.22% Conservative: 70
Best Local Similarity: 26.21% Mismatches: 200
Query Match: 11.91% Indels: 128
DB: 10 Gaps: 23

US-09-712-338-1 (1-1668) x US-09-712-338-1 (1-476)
QY 50 GGGCCCTTCAGGAAGTACACCGCGCTCGTGGTAGAAGACAGTACCCAGAACCCCA 109
Db 22 GlyLeuPheArgSerLeuTyrArgSerValSerMet-----Pro 34
QY 110 CCGGGGTCAAGACTCTTACAAACCGCAAAACAATGTCACCA-----TCGGTACAGG 160
Db 35 ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle 54
QY 161 AACCCGGGCAGAGGGCGTCTGCGAGACTACACCGG-----GGTGCATCAATCTTAC 210
Db 55 GlnLysGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr 74
QY 211 TCTGGATAT-----GTCGACACTCTCCGAGTCCCATACCTTCTTCTGTTCTTCCAA 264
Db 75 AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPheThrPhePro 94
QY 265 GCCAGACATAACCCAGAAACTGCACCTATCACATTTGTGGTGGTGGCCCTCGGAAGC 324
Db 95 AlaGlnIleGlnProGluAspAlaProValIleLeuThrLeuGlnGlyProGlyGly 114
QY 325 GATTCTTTGATCGGTCTCTTCGAAAGAGTTGGGCGCTTGGCCATGTCATTCGACTTTGAT 384
Db 115 SerSerMetPheGlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr 134
QY 385 GACTACATCAACCTCTACTCGTGAACAGAGTCTCCAAATTTACTATTCTCTCCAGCCA 444
Db 135 LeuArgAspArgAspPheProTyrThrThrLeuSerMetLeuTyrIleAspAsnPro 154
QY 445 TTGGGAGTCGGCTTTTCATATAGTGATACGCTTGGTGGTCCATTAACCCCTGTAACTGGG 504
Db 155 ValGlyThrGlyPheSerPheThrAspAspThrHisGly----- 167
QY 505 GTCGTCGAAATTCGAGCTTTGAGGAGTTGAGGGCCGCTTACCAACCAATTTGATGCCACT 564
Db 168 -----Tyr 168
QY 565 CTGATCGATACCAATTTTGGCGCAGAGCGCTTGGGAGATCTCTCAAGAGTTCCTT 624
Db 169 AlaValAsnGluAspAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe 187
QY 625 AGTGGACTACCTAGCTTGGACTCTTAGGGTGAGTCTAAGAGCTTCAAGCTCTATGGAGGAG 684
```

```
Db 188 -----GluIlePheProGluTyrLysAsnAspPheTyrValThrGlyGlu 203
QY 685 AGCTATGAGGCGACTATGGTCTCTCATCTTTCATCATTTTACGACGAGATGAGA 744
Db 204 SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn----- 221
QY 745 ATTGCCAACGGTAGTGTAAATGGTGTTCAGCTTAAATTTCAACTCTCTCGGAATATTAAAC 804
Db 222 -----ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp 237
QY 805 GGCATCATCGAGGCGATCGAGCCCTTACTACCTGCTGAATTCGCTGTGCAACAATACC 864
Db 238 GlyTyrSerAspProGluSerIleIleGlyTyrAlaGluPhe-----Leu 253
QY 865 TACGGTATCAAGGCTGCAACGACCGCTTACAACTACATGAAGTTGCGCAACAATG 924
Db 254 TyrGlnIleGlyLeuLeuAspGlnLysLysTyr-----PheGlnLysGln--- 270
QY 925 CCAATGCTGCCAGGATTGATTTCCACCTGCAACAGACAAAC-----CGCAC 975
Db 271 -----CysHisGluCysIleGluHisIleArgLysGlnAsnTyrPheGluAlaPhe 287
QY 976 GCATTAGTACTAGCGCTGCGCCGACGACCAAC-----ATGTGAGGAGCAAT 1029
Db 288 GluIleLeuAspLysLeuLeuAspGlyAspLeuThrSerAspProSerTyrPheGlnAsn 307
QY 1030 GTTGAGGG-----CCATACCTAGCGCTTCTGCTGCTGCTGATGATATTCGGCAT 1083
Db 308 ValThrGlyCysSerAsnTyrTyrAsnPheLeu----- 318
QY 1084 CCATATGATGACCGCATCTCCGCCAAGTTATTACAAATAATTTCTGGCAAGGACTGTGTC 1143
Db 319 ArgCysThrGluProGluAspGlnLeuTyrTyrValLysPheLeuSerLeuProGluVal 338
QY 1144 ATGAGACCTATCGCGCTCAACATCACTACACCGACTCCAAATATGAGCTCTAC----- 1197
Db 339 ArgGlnAlaIleHisValGly-----AsnGlnThrPheAsnAspGlyThrIleValGluLys 357
QY 1198 TAGCTTTTCCAGAAACAGCGGACTTTGCTGCGCCAACTTCATCGAGACCTCGAGGAG 1257
Db 358 TyrLeuArgGluAspThrValGlnSerValLysProTyrPheThrGluIleMetAsnAsn 377
QY 1258 ATCCTTGCTCTCCCGTGGCTGCTCCCTCATCTATGGCGACGCGCATATCATC----- 1311
Db 378 -----TyrLysValLeuIleTyrAsnGlyGlnLeuAspIleIleValAla 392
QY 1312 -----TGCAACTGTTGCGCGCTGAGGCGGTGAGTGGGAAACTCGCGAG 1341
Db 393 AlaAlaLeuThrGluArgSerLeuMetGlyMetAspTyrLysGlySerGln----- 409
QY 1342 CTGCTGCGAACTACTCCCAAGCGCCGAC-----TTCCGAGC----- 1380
Db 410 -----GluTyrLysLysAlaGluLysValTyrLysIlePheLysSerAspSer 426
QY 1381 -----GCAGGGTACAGCCCTGAAAGTCAACGGCGTCAAGTATGGGAAACTCGCGAG 1434
Db 427 GluValAlaGlyTyr-----IleArgGln 434
QY 1435 TATGTTAATTTCTCTTCTACTCGCTCTATGAGGAGGCGCATGAGTCCCATCTACCGAG 1494
Db 435 AlaGlyAspPheHisGlnValIleIleArgGlyGlyGlyHisIleLeuProTyrAspGln 454
QY 1495 CCCATCGCTCTCTGCAATTTGTTTAAACCGGACTATCTTC-----GGTTGGGAT 1542
Db 455 ProLeuArgAlaPheAspMetIleAsnArgPheIleTyrGlyLysGlyTyrAsp 472
```

RESULT 9

```
US-09-712-338-1 (1-1668) x US-09-712-338-1 (1-1668) (1-476)
; Sequence 164, Application US/0907841
; Publication No. US20020198366A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.

QY 50 GGGCCCTTCCAGGAAGTACACCGCGCTCGCTCGGTAGAGACAGCTACCCAAAGAACCCCA 109
Db 22 GlyLeuPheArgSerLeuTyrArgSerValSerMet-----Pro 34
QY 110 CCGGGGTCAAGACTTTTACACCGCAACAAATGTGTCACCA-----TCGGGTACAAAGG 160
; APPLICANT: Genentech, Inc.
```

```
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,841
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-712-338-1 (1-1668) x US-09-712-338-1 (1-1668) (1-476)
```

```
Alignment Scores:
Pred. No.: 2,92e-21 Length: 476
Score: 360.50 Matches: 141
Percent Similarity: 39.22% Conservative: 70
Best Local Similarity: 26.21% Mismatches: 200
Query Match: 11.91% Indels: 128
DB: 10 Gaps: 23
```

35 ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle 54
 161 AACCGGGGAGAGGGCGTCTCGAGACTACCCCG-----GGTGTCAAAATCTTAC 210
 55 GlnLysGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr 74
 211 TCTGATAT-----GTGACACCTCTCCCGAGTCCCATACCTCTCTCTGTTCTTCGAA 264
 75 AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPhePheTyrPhePro 94
 265 GCCAGACATAACCCAGAACTGACCTATACATTTGTGTGATGTCGAGTGGCCCTGGAAGC 324
 95 AlaGlnIleGlnProGluAspAlaProValLeuThrValLeuThrValGlyProGly 114
 325 GATTTCTTGAUCGCTCTTTCGAAGAGTGGCCCTTGCATGTCAATTCGACTTTTGTAT 384
 115 SerMetPheGlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr 134
 385 GACTACATCAACCCCTCACTCGTGAACGAGGTCTCCAAATTTACTATCTCTGTCGCCAGCA 444
 135 LeuArgAspArgAspPheProTyrThrThrLeuSerMetLeuTyrIleAspAsnPro 154
 445 TTGGAGTCGGCTTTTATATAGTATGATCGGTGATGCTGCTCAATTAACCCCTGTAAGTGG 504
 155 ValGlyThrGlyPheSerPheThrAspAspThrHisGly----- 167
 505 GTCTGCGAAATTCGAGCTTTGAGGAGTTGAGGCGCGGTACCCMAACCATTTGATGCCACT 564
 168 -----Tyr 168
 565 CTGATGCTACTACCAATCTGCGCAGAGCGCGTGGAGATCCTGCAAGATTCCTT 624
 169 AlaValAsnGluAspAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe--- 187
 625 AGTGGACTACCTAGCTTGGACTCTAGGTGTCAGTCTAAGACTTCAGTCTATGAGCGGAG 684
 188 -----GlnIlePheProGluTyrLysAsnAspPheTyrValThrGlyGlu 203
 685 AGCTATGAGGCGACTATGCTGCTGCTGCTTCTCAATTCATTTTACGAGCAGATGAGAGA 744
 204 SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn----- 221
 745 ATTGCCAACGGTATGTTAATGTGTGCTAGCTTAATTTCACTCTCTGCGGAATTTATAC 804
 222 -----ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp 237
 805 GGATCATGACAGAGGCGATCCAGGCCCTTACTACCTGAATTCGCTGTGGAACAATACC 864
 238 GlyTyrSerAspProGluSerIleIleGlyGlyTyrAlaGluPhe-----Leu 253
 865 TACGGTATCAAGGTGTCACAGAGACCGCTCTACAACTACATGAATTTGCCAACCAATG 924
 254 TyrGlnIleGlyLeuLeuAspGluLysLysTyr-----PheGlnLysGln--- 270
 925 CCAAAATGTTCCAGGATTTGATTTCCACCTGCGAAGCAGACAAC-----CGCACC 975
 271 -----CysHisGluCysIleGluHisIleArgLysGlnAsnTyrPheGluAlaPhe 287
 976 GCATTAGCTAGTACGCTCTGCGCGAGACCCCAAC-----ATGTGAGGAGCAAT 1029
 288 GluIleLeuAspLysLeuLeuAspGlyAspLeuThrSerAspProSerTyrPheGlnAsn 307
 1030 GTTGAGGGG-----CCATACCTAGCCCTTTCTGCTGCTGTGTATGATATTCGGCAT 1083
 308 ValThrGlyCysSerAsnTyrTyrAsnPheLeu----- 318
 1084 CCATATGATGACCGCTCTCCCGCAAGTTTATACAAATTTCTGGCAAGGACTCTGTC 1143
 319 ArgCysThrGluProGluAspGlnLeuTyrTyrValLysPheLeuSerLeuProGluVal 338
 1144 ATGACGCTATCGGCGTCAACATCACTACACCCAGTCCCAATATGAGCTCTAC----- 1197
 339 ArgGlnAlaIleHisValGly-----AsnGlnThrPheAsnAspGlyThrIleValGluLys 357

1198 TACCTTTCCAGCAACAGGCGACTTTGTCTGGCCCACTTATCTGAAGACCTCGAGGAG 1257
 358 TyrLeuArgGluAspThrValGlnSerValLysProTyrLeuThrGluIleMetAsnAsn 377
 1258 ATCCTTGTCTCCCGTGTGTCTCCCTCATCTATGCGCAGCGCGGATTTACATC----- 1311
 378 -----TyrLysValLeuIleTyrAsnGlyGlnLeuAspIleIleValAla 392
 1312 -----TGCAACTGTTCCGGCGGTCTGCGGCTTTCC----- 1341
 393 AlaAlaLeuThrGluArgSerLeuMetGlyMetAspTyrLysGlySerGln----- 409
 1342 CTCGCTGGAACTACTCCCAAGCCGCCAG-----TTCGGAAGC----- 1380
 410 -----GluTyrLysLysAlaGluLysLysValTyrLysIlePheLysSerAspSer 426
 1381 -----GCAGGTACACGCCCTCGAAAGTCAACGCGCTGAGTATGGGAACTCCCGAG 1434
 427 GluValAlaGlyTyr-----IleArgGln 434
 1435 TATGTAATTTCTCTTCTCCTCCTATGAGGCGAGCGCATGAAGTCCCATACCTACCGAG 1494
 435 AlaGlyAspPheHisGlnValIleIleArgGlyGlyHisIleLeuProTyrAspGln 454
 1495 CCCATCGCTCCTCCTCAATTTGTTAACCGGACTATCTTC-----GGTGGGAT 1542
 455 ProLeuArgAlaPheAspMetIleAsnArgPheIleTyrGlyLysGlyTyrAsp 472

RESULT 10

US-09-904-011-164
 ; Sequence 164, Application US/09904011
 ; Publication No. US2003000350A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, A.
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth, J.
 ; APPLICANT: Kijavich, Ivar J.
 ; APPLICANT: Mather, Jennie P.
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William, I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: 10466-14
 ; CURRENT APPLICATION NUMBER: US/09/904,011
 ; CURRENT FILING DATE: 2001-07-11
 ; PRIOR APPLICATION NUMBER: 09/665,350
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: PCT/US00/04414
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: US 60/143,048
 ; PRIOR FILING DATE: 1999-07-07
 ; PRIOR APPLICATION NUMBER: US 60/145,698
 ; PRIOR FILING DATE: 1999-07-26
 ; PRIOR APPLICATION NUMBER: US 60/146,222

Db 435 AlaGlyAspPheHisGlnValIleIleArgGlyGlyHisIleLeuProTyrAspGln 454
Qy 1495 CCATCGCTCCGTCATTTTAAACGGGACTATCTTC-----GGTTGGGAT 1542
Db 455 ProLeuArgAlaPheAspMetIleAsnArgPheIleTyrGlyLysGlyTyrAsp 472

RESULT 11
US-09-906-742-164
; Sequence 164, Application US/09906742
; Publication No. US20030023054A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kiljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/906,742
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-742-164
Alignment Scores:
Pred. No.: 2,92e-21 Length: 476
Score: 360.50 Matches: 141
Percent Similarity: 39.22% Conservatives: 70
Best Local Similarity: 26.21% Mismatches: 200
Query Match: 11.91% Indels: 128
DB: 11 Gaps: 23
US-09-712-338-1 (1-1668) x US-09-906-742-164 (1-476)
Qy 50 GGGCCCTTCCAGAAAGTACACGGCGTCGTCGTAGAGACAGCTACCCAGAACCCCA 109
Db 22 GlyLeuPheArgSerLeuTyrArgSerValSerMet-----Pro 34
Qy 110 CCGGGGTCAAGACTCTTACACCGCAACAATGTCAACA-----TCCGGTACAAGG 160
Db 35 ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle 54
Qy 161 AACCCGGGCGAGAGGGCGTCTCGAGACTACCCCG-----GGTGTCAATCTCTAC 210
Db 55 GlnLysGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr 74
Qy 211 TCTGGATAT-----GTCACACCTCTCCGAGTCCCATACCTCTCTCTGTTCTTCGAA 264
Db 75 AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPheThrPhePro 94
Qy 265 GCAGACATAACCCAGAACTGCACCTATCACATTCGTGTGTAATGCTGCTCCCTGGAAGC 324
Db 95 AlaGlnIleGlnProGluAspAlaProValValLeuTrpLeuGlnGlyGlyProGlyGly 114
Qy 325 GATTCTTTGATCGGTCTCTTCGAAGAGTTGGGCGCTTGGCCATGTCATTCGATTTTGAT 384
Db 115 SerSerMetPheGlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr 134
Qy 385 GACTACATCAACCCCTCCTCTCGTGAACGAGGTCTCAATTACTTCTCTGTCGCCACCA 444
Db 135 LeuArgAspArgAspPheProTyrThrThrLeuSerMetLeuTyrIleAspAsnPro 154
Qy 445 TTGGGAGTCGGCTTTTCATATAGTACGGTTGATGGGTCCATTAAACCTGTAACCTGGG 504
Db 155 ValGlyThrGlyPheSerPheThrAspAspThrHisGly----- 167
Qy 505 GTCGTCGAAAATTCGAGCTTTTCAGAGAGTTTCAGGGCCGGTACCCACCACTGTGCGACT 564
Db 168 -----Tyr 168
Qy 565 CTGATCGATACCTACCAATCTTCGCGCAGAGCCGCTTGGGAGATCTTCGAGGATTCCTT 624
Db 169 AlaValAsnGluAspAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe--- 187
Qy 625 AGTGGACTACCTAGCTTGGACTCTAGGGTGCAGTCTAAGGACTTTCAGTCTATGACGGAG 684
Db 188 -----GlnIlePheProGluTyrLysAsnAsnAspPheTyrValThrGlyGlu 203
Qy 685 AGCTATGGAGGCGACTATGGTCCTGCATCTTCAATCATTTTACGACGACAAATCAGAGA 744
Db 204 SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn----- 221
Qy 745 ATTGCCAACGGTAGTGTAAATGGTCTCAGCTTAATTTCACTCTCTGGGATTTAATTAAC 804
Db 222 -----ProValArgGluValLysIleAsnLeuAsnGlyIleAlaIleGlyAsp 237

Qy 50 GGGCCCTTCCAGGAAGTACACCGCGTCGGTCGGTAGAGACAGTCACCCAGACCCCA 109

QY	1144	ATGGAGCGTATCGGCGTCAACATCACTACACCGAGTCCAAATAATGACGCTAC-----	1197
Db	339	ArgGlnAlaIleHisValGly---AsnGlnThrPheAsnAspGlyThrIleValGluLys	357
QY	1198	TAGCGTTTCAGCAAAACAGCGCACTTTGTCTGCGCCCAACTTCATCGAAGACCTCGAGGAG	1257
Db	358	TyrLeuArgGluAspThrValGlnSerValLysProTyrLeuThrGluIleMetAsnAsn	377
QY	1258	ATCCTTGCTCTCCCGTGGCGTGTCTCCCTCATCTATGGCGACGCCGATTACATC-----	1311
Db	378	-----TyrLysValLeuIleIleValGly---TyrLysValLeuIleIleValAla	392
QY	1312	-----TGCAACTGGTTCGCGCTCAGCGCGTTCCTTC	1341
Db	393	AlaAlaLeuThrGluArgSerLeuMetGlyMetAspTrpLysGlySerGln-----	409
QY	1342	CTGCTCGCAACTACTCCCAAGCGGCCAG-----TTCCGAGC-----	1380
Db	410	-----GluTyrLysLysAlaGluLysValTrpLysIlePheLysSerAspSer	425
QY	1381	-----CGAGGTACACGCCCTGAAAGTCAACCGCTCGAGTATGGGAACTCGCGAG	1434
Db	427	GluValAlaGlyTyr-----IleArgGln	434
QY	1435	TATGGTAATTTCTCCTTCATTCGGTCTATGAGCAGCGCATGAAGTCCCATCTACTACAG	1494
Db	435	AlaGlyAspPheHisGlnValIleIleArgGlyGlyGlyHisIleLeuProTyrAspGln	454
QY	1495	CCATCGCCTCCCTGCATTTCTTTAAACGGGACTATCTTC-----GGTTGGGAT	1542
Db	455	ProLeuArgAlaPheAspMetIleAsnArgPheIleTyrGlyLysGlyTyrAsp	472
RESULT 13			
US-09-907-613-164			
; Sequence 164, Application US/09907613			
; Publication No. US20030027145A1			
; GENERAL INFORMATION:			
; APPLICANT: Genentech, Inc.			
; APPLICANT: Ashkenazi, Avi			
; APPLICANT: Botstein, David			
; APPLICANT: Desnoyers, Luc			
; APPLICANT: Eaton, Dan L.			
; APPLICANT: Ferrara, Napoleone			
; APPLICANT: Filvaroff, Ellen			
; APPLICANT: Fong, Sherman			
; APPLICANT: Gao, Wei-Qiang			
; APPLICANT: Gerber, Hanspeter			
; APPLICANT: Gerritsen, Mary E.			
; APPLICANT: Goddard, A.			
; APPLICANT: Godowski, Paul J.			
; APPLICANT: Grimaldi, Christopher J.			
; APPLICANT: Gurney, Austin L.			
; APPLICANT: Hillan, Kenneth, J.			
; APPLICANT: Kljavin, Ivar J.			
; APPLICANT: Mather, Jennie P.			
; APPLICANT: Pan, James			
; APPLICANT: Paoni, Nicholas F.			
; APPLICANT: Roy, Margaret Ann			
; APPLICANT: Stewart, Timothy A.			
; APPLICANT: Tumas, Daniel			
; APPLICANT: Williams, P. Mickey			
; APPLICANT: Wood, William, I.			
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic			
; FILE REFERENCE: 10466-14			
; CURRENT APPLICATION NUMBER: US/09/907,613			
; CURRENT FILING DATE: 2001-07-17			
; PRIORITY APPLICATION NUMBER: PCT/US00/04414			
; PRIORITY FILING DATE: 2000-02-22			
; PRIORITY APPLICATION NUMBER: US 60/143,048			
; PRIORITY FILING DATE: 1999-07-07			
; PRIORITY APPLICATION NUMBER: US 60/145,698			

Db 135 LeuArgAspArgAspPheProTrpThrThrLeuSerMetLeuTyrlleAspAsnPro 154

427 GluValAlaGlyTyr-----IleArgGln 434

```

; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-942-164

Alignment Scores:
Pred. No.: 2,92e-21 Length: 476
Score: 360.50 Matches: 141
Percent Similarity: 39.22% Conservative: 70
Best local Similarity: 26.21% Mismatches: 200
Query Match: 11.91% Indels: 128
DB: 11 Gaps: 23

US-09-712-338-1 (1-1668) x US-09-907-942-164 (1-476)
QY 50 GGGCCCTTCCAGGAAGTACACCGCGCTCGTGGTAGAGACGCTACCCAGAACCCCA 109
Db 22 GlyLeuPheArgSerLeuTyrArgSerValSerMet-----Pro 34
QY 110 CCGGGGTCAAGACTCTTACACCGCAACAAATGTCCACCA-----TCCGGTACAGG 160
Db 35 ProLysGlyAspSerGlyGlnProLeuPheLeuThrProTyrIleGluAlaGlyLysIle 54
QY 161 AACCCGGGCGAGAGGCGTCTCGAGACTACCCCG-----GGTCTCAATCCTAC 210
Db 55 GlnLysGlyArgGluLeuSerLeuValGlyProPheProGlyLeuAsnMetLysSerTyr 74
QY 211 TCTGGATAT-----GTCACACCTCTCCGAGTCCCATACCTTCTTCTGTTCTTCTGAA 264
Db 75 AlaGlyPheLeuThrValAsnLysThrTyrAsnSerAsnLeuPhePhePhePro 94
QY 265 GCAGACATAACCCAGAACTGCACCTATCACATTTGTGTGAATGGTGGCCCTCGAAGC 324
Db 95 AlaGlnIleGlnProGluAspAlaProValValLeuThrLeuGlnGlyProGlyGly 114
QY 325 GATTCTTTGATCGTCTCTTCGAGAGTTGGGCGCTTCGCATGTCAATTCGACTTTTGAT 384
Db 115 SerSerMetPheGlyLeuPheValGluHisGlyProTyrValValThrSerAsnMetThr 134
QY 385 GACTATCATCAACCTCACTCTCGGAACGAGGTCTCCAAATTTACTATTCTGTCCAGCCA 444
Db 135 LeuArgAspArgAspPheProTyrThrThrLeuSerMetLeuTyrIleAspAsnPro 154
QY 445 TTGGGAGTCGGCTTTTCATATAGTATACGGTTGATGGGTCCATTAACTGTAACTGGG 504
Db 155 ValGlyThrGlyPheSerPheThrAspAspThrHisGly-----167
QY 505 GTCGTCGAAAATTGAGCTTTGACAGGATTCAGGGCCGTTACCCACCATTTGTCGCACT 564
Db 168 -----Tyr 168
QY 565 CTGATCGATACCTACCAATCTTCCGCGAGAGCCGCTTGGGAGATCTTGGCAAGATTCCTT 624
Db 169 AlaValAsnGluAspAspValAlaArgAspLeuTyrSerAlaLeuIleGlnPhePhe--- 187
QY 625 AGTGGACTACTAGCTTGGACTCTAGGGTGCAGGTCTAAGGACTTTCAGTCTATGACCGAG 684
Db 188 -----GlnIlePheProGluTyrLysAsnAsnAspPheTyrValThrGlyGlu 203
QY 685 AGCTATGAGGGCACTATGCTCTGCAATCTTCAATCATTTTACGACGACGAATGAGAGA 744
Db 204 SerTyrAlaGlyLysTyrValProAlaIleAlaHisLeuIleHisSerLeuAsn----- 221
QY 745 ATTGCCAACGGTAGTGTATTGCTTACGCTTATTTCACTCTCTGGGATTTATTAAAC 804
Db 222 -----ProValArgGluValLysIleAsnGlyIleAlaIleGlyAsp 237
; PRIOR FILING DATE: 1999-12-20

RESULT 14
US-09-907-942-164
; Sequence 164, Application US/09907942
; Publication No. US20030027145A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mackey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,942
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
```

[illegible]

Search completed: November 21, 2003, 17:49:19
Job time : 85.5 secs

[illegible]

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 17:37:34 ; Search time 29 Seconds
(without alignments)
4867.208 Million cell updates/sec

Title: US-09-712-338-1
Perfect score: 3027
Sequence: 1 atcggtggctacgaattctt.....ccadgtgtgtatgcatag 1668

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DRV=xlp
-Q=/cg2_1/USPTO.spool_p/US09712338/runat 17112003 170149 10565/app_query.fasta_1.1863
-DB=Issued Patents AA -Qfmt=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @CGN 1.1.31 @runat 17112003 170149 10565 -NCPUS=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBILOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:**
1: /cg2_6/prodata/2/iaa/5A.COMB.pep.*
2: /cg2_6/prodata/2/iaa/5B.COMB.pep.*
3: /cg2_6/prodata/2/iaa/6A.COMB.pep.*
4: /cg2_6/prodata/2/iaa/6B.COMB.pep.*
5: /cg2_6/prodata/2/iaa/6C.COMB.pep.*
6: /cg2_6/prodata/2/iaa/6D.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2975	98.3	554	3	US-08-943-714-2
2	894	25.5	423	3	US-08-943-714-9
3	630	20.8	557	1	US-08-309-341-2
4	630	20.8	557	1	US-08-608-267-2
5	630	20.8	557	1	US-08-608-452-2
6	630	20.8	557	1	US-08-608-224-2
7	630	20.8	557	2	US-08-967-149-2
8	623	20.6	557	1	US-08-309-341-4
9	623	20.6	557	1	US-08-608-267-4
10	623	20.6	557	1	US-08-608-452-4
11	623	20.6	557	1	US-08-608-224-4
12	623	20.6	557	2	US-08-967-149-4

13	568.5	18.8	491	1	US-09-640-305-4	Sequence 4, Appli
14	568.5	18.8	491	1	US-08-360-673-4	Sequence 4, Appli
15	523.5	17.3	532	3	US-08-899-324-33	Sequence 33, Appl
16	523.5	17.3	532	3	US-08-329-892B-33	Sequence 33, Appl
17	513.5	17.0	421	2	US-08-807-263-4	Sequence 4, Appli
18	361.5	11.9	476	2	US-08-828-488-3	Sequence 3, Appli
19	361.5	11.9	476	4	US-09-299-689A-3	Sequence 3, Appli
20	349	11.5	477	2	US-08-828-488-1	Sequence 1, Appli
21	349	11.5	477	4	US-09-299-689A-1	Sequence 1, Appli
22	337	11.1	480	2	US-08-828-488-8	Sequence 8, Appli
23	337	11.1	480	4	US-09-299-689A-8	Sequence 8, Appli
24	337	11.1	480	4	US-09-702-705-336	Sequence 336, App
25	337	11.1	480	4	US-08-828-488-7	Sequence 336, App
26	312.5	10.3	471	2	US-08-828-488-7	Sequence 7, Appli
27	312.5	10.3	471	4	US-09-299-689A-7	Sequence 7, Appli
28	267	8.8	523	3	US-08-943-714-11	Sequence 11, Appl
29	246.5	8.1	446	1	US-08-665-966-10	Sequence 10, Appl
30	246.5	8.1	446	3	US-09-041-780-10	Sequence 10, Appl
31	208	6.9	530	3	US-08-943-714-12	Sequence 12, Appl
32	204.5	6.8	481	3	US-08-943-714-10	Sequence 10, Appl
33	193	6.4	35	3	US-08-943-714-5	Sequence 5, Appli
34	172	5.7	179	1	US-08-665-966-8	Sequence 8, Appli
35	172	5.7	179	3	US-09-041-780-8	Sequence 8, Appli
36	169.5	5.6	351	2	US-08-828-488-5	Sequence 5, Appli
37	169.5	5.6	351	4	US-09-299-689A-5	Sequence 5, Appli
38	163.5	5.4	878	4	US-09-556-706B-2	Sequence 2, Appli
39	161	5.3	907	3	US-08-783-774-2	Sequence 2, Appli
40	161	5.3	907	4	US-09-328-599A-1	Sequence 1, Appli
41	161	5.3	907	5	PCT-US95-04611A-19	Sequence 19, Appl
42	156	5.2	1048	4	US-09-171-699-10	Sequence 10, Appl
43	155.5	5.1	1911	4	US-09-854-856-64	Sequence 64, Appl
44	155.5	5.1	1939	4	US-09-854-856-48	Sequence 48, Appl
45	155.5	5.1	1971	4	US-09-854-856-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-08-943-714-2
; Sequence 2, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berk, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dammann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,714
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4990.200-US
; TELECOMMUNICATION INFORMATION:

Mon Nov 24 13:41:03 2003

TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
SEQUENCE CHARACTERISTICS:
LENGTH: 554 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-943-714-2

Alignment Scores:
Pred. No.: 2,346-275 Length: 554
Score: 2975.00 Matches: 554
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.28% Indels: 0
DB: 3 Gaps: 0

US-09-712-338-1 (1-1668) x US-08-943-714-2 (1-554)

QY 1 ATGCGTGGCTACGAATTTCTCTCAGTGTACCCCTTGGTGGAGCGAGTGGCCCTTCCA 60
DB 1 MetArgGlyTyrGluPheLeuSerValLeuProLeuValAlaAlaSerTrpAlaLeuPro 20
QY 61 GGAAGTACACCGGCTCGGTAGAACAGACAGTACCCAGAACCCACCGGGGTCAAG 120
DB 21 GlySerThrProAlaSerValGlyArgArgGlnLeuProGlyAsnProThrGlyValLys 40
QY 121 ACTCTTACAAACGCAACATTCACATCCGGTACAGGAACCCGGGCGAGGCGTC 180
DB 41 ThrLeuThrAlaAlaAsnValThrIleArgTyrLysGluProGlyAlaGluGlyVal 60
QY 181 TGGAGACATACCCCGGTGTCAATCTCTATCTGGATATGTGACACCTCTCCGAGTCC 240
DB 61 CysGluThrThrProGlyValLysSerTyrSerGlyTyrValAspThrSerProGluSer 80
QY 241 CATACCTTCTTGGTCTTCGAGCCAGACATACCCAGAACCCGACTATCACATTG 300
DB 81 HisThrPheThrPhePheGluAlaArgHisAsnProGluThrAlaProIleThrLeu 100
QY 301 TGGTTCGAATGCTGCGCTGGAAGCGATTCTTTCATCGGTCTCTTCAAGAGATTGGCCCT 360
DB 101 TrpLeuAsnGlyGlyProGlySerAspSerLeuIleGlyLeuPheGluLeuGlyPro 120
QY 361 TGGCATGTCAATTCGACTTTTGTATGACTACATCAACCCCTCAGTGGAGCGAGTCTCC 420
DB 121 CysHisValAsnSerThrPheAspTyrIleAsnProHisSerTrpAsnGluValSer 140
QY 421 AATTACTATCTCTGCTCCGACCATTCGGAGTGGCTTTTCATATAGTATGATACGTTGAT 480
DB 141 AsnLeuLeuPheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAsp 160
QY 481 GGGTCCATTAAACCCCTTAACCTGGGGTGTCTGAAATTCGAGCTTTGCGAGGATTTCAGGC 540
DB 161 GlySerIleAsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGly 180
QY 541 CGGTACCCAAACCATTTGATGCACTCTGATGATACCTACCAATCTTGGCGAGAGCGCT 600
DB 181 ArgTyrProThrIleAspAlaThrLeuIleAspThrAsnLeuAlaAlaGluAlaAla 200
QY 601 TGGAGATCTGCGAGGATTCCTTAGTGGACTACCTAGCTTGGACTCTAGGCTGAGTCT 660
DB 201 TrpGluIleLeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSer 220
QY 661 AAGGACTTCAGTCTATGGACGAGAGCTATGAGGGCACTATGCTCTGCAATTTCTCAAT 720
DB 221 LysAspPheSerLeuTrpThrGluSerTyrGlyGlyHisTyrGlyProAlaPheAsn 240
QY 721 CATTTTACGACAGAGATTCGACAGATTCGACAGGAGTGTAAATGGTGTTCAGCTTAAT 780

DB 241 HisPheTyrGluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsn 260
QY 781 TTCAACTCTCTGGGAATTTAAGCGCATCATCCAGCAGGCGATCCAGGCCCTTACTTAC 840
DB 261 PheAsnSerLeuGlyIleIleAsnGlyIleIleAspGluAlaIleGlnAlaProTyrTyr 280
QY 841 CCTGAATTCGTGTGAACAATACCTACGGTATCAAGCTGTCAACGAGACCGCTCTACAC 900
DB 281 ProGluPheAlaValAsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsn 300
QY 901 TACATGAAGTTTGGCCAAACCAATGCCAAATGGTTCGCCAGGATTCGATTCACCTCGCAA 960
DB 301 TyrMetLysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSerThrCysLys 320
QY 961 CAGACAAACCGCACCGCATTTAGCTGACTACCCCTCTCGCGGAGCCACCAACATGTGC 1020
DB 321 GlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCys 340
QY 1021 AGGACAAATGTGGGGGCCATACCTACGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
DB 341 ArgAspAsnValGluGlyProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArg 360
QY 1081 CATCCATATGATGACCCGACTCCGCAAGTTATTAACAATTTCTGGCAAGACGCT 1140
DB 361 HisProTyrAspAspProThrProProSerTyrTyrAsnLysPheLeuAlaLysAspSer 380
QY 1141 GTCATGGAGCTATCGGCTGCAACATCACTACACCCAGTCCCAATTAATGACCTCTACTAC 1200
DB 381 ValMetAspAlaIleGlyValAsnIleAsnTyrThrGlnSerAsnAsnAspValTyrTyr 400
QY 1201 GCTTTCCAGCAACAGGCGACTTGTCTGGCCCAACTTCATCATCAAGACCTCGAGGAGATC 1260
DB 401 AlaPheGlnGlnThrGlyAspPheValTyrProAsnPheIleGluAspLeuGluIle 420
QY 1261 CTTGTCTCTCCGCTGCTGCTCTCCCTCATCTATGCGGACGCCGATCTACTCTCACTCG 1320
DB 421 LeuAlaLeuProValArgValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTrp 440
QY 1321 TTCGGGGTTCAGGCGCTTCCCTCGCTGCGAATCTCCCAAGCCGCCAGTTCCGAGC 1380
DB 441 PheGlyGlyGlnAlaValSerLeuAlaAlaAsnTyrSerGlnAlaAlaGlnPheArgSer 460
QY 1381 GCAGGTTCACAGCCCTCGAAAGTCAACGCGCTCGAGTATGGGGAACCTCCGAGTATGGT 1440
DB 461 AlaGlyTyrThrProLeuLysValAsnGlyValGluTyrGlyGluThrArgGluTyrGly 480
QY 1441 AATTTCTCTCTCACTCCGCTATGAGGCGAGCCATGAGTCCCATCTACTACGAGCCATC 1500
DB 481 AsnPheSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProIle 500
QY 1501 GCCTCCCTGCAATTTGTATTACCGGACTATCTTCGTTGGGATATCGAGAGGCGCAGAAG 1560
DB 501 AlaSerLeuGlnLeuPheAsnArgThrIlePheGlyTrpAspIleAlaGluGlyGlnLys 520
QY 1561 AAGATCTGCGCCAGCTACAAAGACGAATGGAACCGCTACAGCTACGATACAGCTCGTCC 1620
DB 521 LysIleTrpProSerTyrLysThrAsnGlyThrAlaThrAlaThrHisThrGlnSerSer 540
QY 1621 GTGCGCTGCTACCGCTACCGATCTCCAGTCTCGTATG 1662
DB 541 ValProLeuProThrAlaThrSerMetSerValGlyMet 554

RESULT 2
US-08-943-714-9
; Sequence 9, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berk, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik

; APPLICANT: Dambmann, Claus
 ; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 61875780 No. 61875780disk of No. 61875780th America, Inc.
 ; STREET: 405 Lexington Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10174
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/943,714
 ; FILING DATE: 03-OCT-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lambiris, Elias J
 ; REGISTRATION NUMBER: 33,728
 ; REFERENCE/DOCKET NUMBER: 4990.200-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 423 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-943-714-9

Alignment Scores:
 Pred. No.: 1,116-76 Length: 423
 Score: 894.00 Matches: 186
 Percent Similarity: 54.31% Conservative: 66
 Best Local Similarity: 40.09% Mismatches: 160
 Query Match: 29.53% Indels: 52
 DB: 3 Gaps: 8

US-09-712-338-1 (1-1668) x US-08-943-714-9 (1-423)

QY 175 GCGCTGTGCGAGACTACCCGGGTGCAAACTCTACTCTGGATATCTGCACACTCTCCC 234
 Db 6 GlylleCysGluThrThrProGlyValAsnGlnTyrSerGlyTyrLeuSerValGlySer 25
 QY 235 GAGTCCCATACCTTCTTGGTCTTTCGAAGCCAGACATAACCCAGAACTGCACCTATC 294
 Db 26 AsnMetAsnMetTyrPheTyrPheGluAlaAsgAsnAsnProGlnGlnAlaProLeu 45
 QY 295 ACATTGTGTTCAANTGGTGGCCCTGGAAGCGGATCTTTGATCGTCTCTTCGAAGATTG 354
 Db 46 AlaAlaTyrPheAsnGlyGlyProGlyCysSerSerMetIleGlyLeuPheGlnGluAsn 65
 QY 355 GGCCCTTCCCAT--GTCAATTTCGACTTTTGNATGACTACATCAACCTCATCTGPGGAAC 411
 Db 66 GlyProCysHisPheValAsnGlyAspSerThrProSerLeuAsnGluAsnSerTyrPsn 85
 QY 412 GAGGTCTCAATTACTATCTGTCCCGCATTTGGAGTCGGGCTTTTTCATATAGTCAT 471
 Db 86 AsnTyrAlaAsnMetIleTyrIleAspGlnProIleGlyValGlyPheSerTyrGly--- 104
 QY 472 ACGGTGTAGTGGTCCATTAACTGTAACCTGTAACTGGGTGCGGAAATTCGAGCTTTCGAGGA 531
 Db 104 ----- 104
 QY 532 GTTCAGGCGCGGTACCCAAACCATTCATGCCACTCTGATCGATACCTACCAATCTTGGCCGA 591
 Db 105 -----ThrAspAspValThr-----SerThrValThrAlaAla 115

QY 592 GAGCCCGCTTGGGAGATCTCTGCAAGGATTCCTTAGTGGACTACCTAGCTTGGACTCTAGG 651
 Db 116 ProTyrValTyrPsnLeuLeuGlnAlaPheTyrAlaGlnArgProGluTyrGluSerArg 135
 QY 652 GTGCAGCTAAGGACTTCACTATGACGAGAGCTATGAGGGGCACTATGGTCCCTGCA 711
 Db 136 -----AspPheAlaIlePheThrGluSerTyrGlyGlyHisTyrGlyProGlu 151
 QY 712 TTCTTCAATCATTTTACGAGCAGAATGAGAGATTGCCAACGGTACTGTGTAATGGTGT 771
 Db 152 PheAlaSerTyrIleGluGlnGlnAsnAlaAlaIleLysAlaGlySerValThrGlyGln 171
 QY 772 CAGCTTAATTTCAACTCTCTGGGAATTTAATTAACGGCATCATCGACGAGGCGATCCAGGCC 831
 Db 172 AsnValAsnIleValAlaLeuGlyValAsnAsnGlyTyrPheAspSerThrIleGlnGlu 191
 QY 832 CCTTACTACCTGAATTCGCTGTGAACATACCTACGGTATCAAGGCTGTCAACGAGACC 891
 Db 192 LysAlaTyrIleAspPheSerTyrAsnAsnSerTyrGlnGlnIleIleAspSerSerThr 211
 QY 892 GTCTACAACCTACATGAAGTTTGCAACCAATGCCAAATGGTTCGCCAGGATTTGATTCC 951
 Db 212 ArgAspSerLeuLeuAspAlaTyrAsn-----AsnGlnCysLeuProAlaLeuGln 228
 QY 952 ACCTGCAACAGACAAACCCGACCTAGCTAGCTAGCGCCTCTCGCGCGAAGCCACC 1011
 Db 229 GlnCysSerGlnSerGlySerThr-----SerAspCysThrAsnAlaAsp 243
 QY 1012 AACATGTGCGAGGACAAATGTCAGGGGCGCATACTACGCTTCTGCTGGTGGTGTGTAT 1071
 Db 244 SerValCysTyrGlnAsnIleGluGlyProIleSerSerSerGlyAspPheAspValTyr 263
 QY 1072 GATATTGGCATCTCATATGATGACCCGACTCCGCGCAAGTTATTACAACTTCTGGCA 1131
 Db 264 AspIleArgGluProSerAsnAspProTyrProIleSerSerThrTyrLeuSer 283
 QY 1132 AAGGACTGTGTCATGAGCGCTATCGCGGTCACATCACTACACCCAGCTCCAATATGAC 1191
 Db 284 AspProThrValValIleAlaIleGlyAlaArgThrAsnTyrGlnGluCysProAsnGly 303
 QY 1192 GTCTACTAGCTTCCAGCAACAGCGGACTTTGTGCGCCCACTTCATCGAAGACCTC 1251
 Db 304 ProTyrAsnLysPheAlaSerThrGlyAspAsnPro---ArgSerPheLeuSerThrLeu 322
 QY 1252 GAGGAGATCCTTGTCTCTCCCGTGGGTGTCTCCCTCATCTATGGCGAGCCGATACATC 1311
 Db 323 SerSerValValGlnSerGlyIleAsnValLeuValTyrAlaGlyAspAlaAspTyrIle 342
 QY 1312 TGCAACTGGTTCGGCGGTGAGCCGCTTTCCTCTGCTGCGAACTACTCCCAAGCCGCCAG 1371
 Db 343 CysAsnTyrLeuGlyAsnTyrGluValAlaAsnAlaValAspPheProGlyAsnAlaGln 362
 QY 1372 TTCGAACCGGAGGTACACGCCCTTGAAAGTCAACGGCGTCGAGTATGGGAACTCGC 1431
 Db 363 PheSerAlaLeuAspLeuAlaProTyrThrValAsnGlyValGluLysGlyGlnPheLys 382
 QY 1432 GAGTATGTAATTTCT 1491
 Db 383 ThrValAspAsnPheSerPheLeuLysValTyrGlyAlaGlyHisGluValProTyrTyr 402
 QY 1492 CAGCCCATCGCCTCCCTGCAATTTTAAACCGGACTATCTTCGGTGGGATATGCGAGAG 1551
 Db 403 GlnProAspThrAlaLeuGlnAlaPheLysGlnIleIle----- 415
 QY 1552 GGCCAGAGAGAG 1563
 Db 416 ---GlnLysLys 418

RESULT 3
 US-08-309-341-2
 ; Sequence 2, Application US/0809341
 ; Patent No. 5594119

Alignment Scores:		
Pred. No.:	2,13e-51	557
Score:	630.00	154
Percent Similarity:	48.45%	Matches: 81
Best Local Similarity:	31.75%	Mismatches: 174
Query Match:	20.81%	Indels: 76
DR:	1	Gaps: 17

QY	133	GCAACAATGTCACCATCCGGTCAAGGAACCGGGGCAGAGGCGTGCTGGAGACTACC	192
Db	132	AlatyAspIeuArgVallyLysThrAspProGlySerLeuGlylle-----Asp	148
QY	193	CCGGGTGTCAAATCCTCTCGGATATGTGGACACCTCCCC---GAGTCCCATCACCTTC	249
Db	149	ProGlyVallySginTyThrGlyLeuAspAsnGluAsnAspLysHiLeuPhe	168
QY	250	TTCGTGGTTCTCGAAGCAGCACATAACCCAGAACTGCACCTATCACATCTGTGGTTGAAT	309
Db	169	TyrTrpPhePheGluSerArgAsnAspProGluAsnAspProvAluValIeuTrpIeuAsn	188
QY	310	GGTGGCCCTGGAAGCGATTCTTTGATCGGTCTCTTCGAAGAGTTGGGCCCTTGCCATGTC	369
Db	189	GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerile	208
QY	370	AAMTCGACTTTTGATGACTACATCAACCCCTCACTCGTGGAACGAGGTCTCCAATTACTA	429
Db	209	AsnLysLysIleGlnProValTyrAsnAspTyrAlatrpAsnSerAsnAlaSerValille	228
QY	430	TTCCTGTCCCAGCCATTGGGAGTCGGCTTTTCATATAGTAGTACGGTTGATGGGTCCCAAT	489
Db	229	PheLeuAspGlnProValAsnValGlyIyrSerIyrSerAsnSerAla-----	244

Db 552 LeuGlyGlyGluTrp 556

RESULT 4

US-08-608-267-2

; Sequence 2, Application US/08608267

; Patent No. 5688663

; GENERAL INFORMATION:

; APPLICANT: Yaver, Debbie Sue

; APPLICANT: Thompson, Sheryl Ann

; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 5688663o No. 5688663disk of No. 5688663th America, Inc.

; STREET: 405 Lexington Avenue, Suite 6400

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/608,267

; FILING DATE: 28-FEB-1996

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/309,341

; FILING DATE: 20-SEP-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lowney, Karen A.

; REGISTRATION NUMBER: 31,274

; REFERENCE/DOCKET NUMBER: 4247,000-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 867 0123

; TELEFAX: 212 867 0298

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 557 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: Aspergillus Niger

US-08-608-267-2

Alignment Scores:

Pred. No.:	2.13e-51	Length:	557
Score:	630.00	Matches:	154
Percent Similarity:	48.45%	Conservative:	81
Best Local Similarity:	31.75%	Mismatches:	174
Query Match:	20.81%	Indels:	76
DB:	1	Gaps:	17

US-09-712-338-1 (1-1668) x US-08-608-267-2 (1-557)

QY 133 GCAACAATGTCCATCCGGTACAGGAACCGGGGAGAGGGCGTCTGGAGACTACC 192

Db 132 AlaTyrAspLeuArgVallylsThrAspProGlySerLeuGlyle-----Asp 148

QY 193 CCGGTGTCAATCTCTCTGATATGTCGACACTCTCC-----GAGTCCCATACCTTC 249

Db 149 ProGlyVallylsGlnTyThrGlyTyrLeuAspAspAsnGluAsnAspHisLeuPhe 168

QY 250 TTCTGTCTTCAGAGCCAGACATACCCAGAACTGACCTATCACATTTGGTTGAAT 309

Db 169 TyrTrpPheGluSerArgAsnAspProGluAsnAspProValValLeuTrpLeuAsn 188

QY 310 GTGGCCCTCGAAGCGATTCTTTGATCGTCTCTCGAAGAGTTGGGCCCTTGGCCATGTC 369

Db 189 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerile 208

QY 370 AATTCGACTTTTGATGACTACATCAACCTCGTGAACGAGGTCTCCAAATTTACTA 429

Db 209 AsnLysLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValle 228

QY 430 TTCCTGTCCAGCCATCGGAGTGGCTTTTCATATAGTATGATGACGGTTGATGGTCCATT 489

Db 229 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla----- 244

QY 490 AACCTGTAACTGGGGTCTGCGAAAATTCGAGCTTTGAGAGTTCGAGGCGCGGTACCCA 549

Db 244 ----- 244

QY 550 ACCATTGATGCCACTCTGATCATACCAATCTTGCCGAGAGGCGCGTTGGGAGATC 609

Db 245 -----ValSerAspThrValAlaAlaGlyLysAspValTyrAlaLeu 258

QY 610 CTGCAAGGATTCTTAGTGAGCTACCTAGCTTGAGCTCTAGGGTGCACTTAAGGACTTC 669

Db 259 LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe 274

QY 670 AGTCTATGACGGAGAGCTATGAGAGGCATGATGCTCTGCATTCCTCAATCATTTTAC 729

Db 275 HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu 294

QY 730 GAGCAGATGAGAAATTCGCAACGGTAGTGTAAATGGTGTGAGCTTAATTTCAACTCT 789

Db 295 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 305

QY 790 CTGGGAATTATTAAACGGCATCATCGAGAGCGCATCCAGCCCTTACTACCTGGAATTC 849

Db 306 ValLeuIleGlyAsnGlyLeuThrAspGlyTyrThrGlnTyrGluTyrTyrArgProMet 325

QY 850 GCTGTGAACAATACCTAGGTATCAAGGCTGTCAAGAGGCTGTCAAGAGACCTCTACAAC 909

Db 326 AlaCysGlyAspGlyGlyTyrProAlaValLeuAspGluSerSerCysGlnSerMetAsp 345

QY 910 TTGCGCAACCAATGCCAATGTTGCCAGGATTTGATTTCCACCTTCCAAACAGACAAAC 969

Db 346 -----AsnAlaLeuProArg--CysGlnSerMetIleGluSerCysTyrSerSerGlu 362

QY 970 CGCACCGCATTAGCTAGCTAGCCCTCTGCGCGAAGCCACCAACATGTGAGGAGCAAT 1029

Db 363 SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAsnAla 377

QY 1030 GTTGAGGGCCCATACTACGCCCTTGTGCTGCTGTGTGTATGATATTCGGCATCCATAT 1089

Db 378 LeuLeuAlaProTyrGlnArg--ThrGlyGlnAsnValTyrAspValArgGlyLysCys 396

QY 1090 GATGACCCG-----ACTCGCCAAAGTTATTACAAACAAATTTCTGGCAAG 1134

Db 397 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 416

QY 1135 GACTCTGTCATGCGAGCTATCGGCTCAACATCAAC---TACACCCAGTCCCAATATGAC 1191

Db 417 ProGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 436

QY 1192 GTCTACTACGCTTTCAGCAACAGCGACTTTGTCTGGCCC-----AATTCTATCGAA 1245

Db 437 IleAsnArgAsnPheLeuPheHisGlyAspTrpMetLysProTyrHisArgLeuValPro 456

QY 1246 GACTCGAGGAGATCTTGTCTCTCCCGTGTCTCTCCCTCATCTAT---GGCGACGCC 1302

Db 457 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 471

QY 1303 GATTACATCTCAACTGTTTCGGCGGTGAGCGCTTCCCTCGCTCGCAATCTATCCCAA 1362

Db 472 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTyrProGly 491

QY 1363 GCCGCCCATGTTCCGAAGCGAGGTACACGCCCTCTGAAGTGC-----AAC 1407

Db 492 GlnAlaGluTyrAlaSerAlaGluLeuAspLeuValIleValAspAsnGluHisThr 511

```

1408 GCGCTGCGATGATGGGAACACTCGCGAGTAGTGTAATTTCCTTCACTCGCGTCTATGAG 1467
||||| :::: ||||| :::: ||||| :::: ||||| :::: ||||| :::: ||||| :::: |||||
512 GlyLysLysIleGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyGly 531
1468 GCAGGCCCATGAGTCCCATCTACTACGACCCTCGCCTCCCTGCATTTGTAAACCGGACT 1527
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
532 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPheAsnArgTrp 551
1528 ATCTTCGGT---TGG 1539
||||| :::: ||||| :::: ||||| :::: ||||| :::: ||||| :::: ||||| :::: |||||
552 LeuGlyGlyGluTrp 556

RESULT 5
US-08-608-452-2
; Sequence 2, Application US/08608452
; Patent No. 5693510
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5693510c No. 5693510disk of No. 5693510th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,452
; FILING DATE: 28-FEB-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus Niger
US-08-608-452-2

Alignment Scores:
Pred. No.: 2,13e-51 Length: 557
Score: 630.00 Matches: 154
Percent Similarity: 48.45% Conservative: 81
Best Local Similarity: 31.75% Mismatches: 174
Query Match: 20.81% Indels: 76
DB: 1 Gaps: 17

US-09-712-338-1 (1-1668) x US-08-608-452-2 (1-557)
QY 133 GCAAAACAATGTCACCATCCGGTACAAAGAACCCCGGGCGAGCGGCTCTCGAGACTACC 192
132 AGTatv-asoLeuAroGValLysThrAspProGlySerLeuGlyIle-----Asp 148

```


LENGTH: 557 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Aspergillus Niger
US-08-967-149-2

Alignment Scores:	2.13e-51	Length:	557
Pred. No.:	630.00	Matches:	154
Score:	38.45%	Conservative:	81
Percent Similarity:	48.75%	Mismatches:	174
Best Local Similarity:	38.1%	Indels:	76
Query Match:	20.81%	Gaps:	17
DB:	2		

713-338-1 (1-1668) x 11S-08-967-149-2 (1-557)

133	QY	133	GC	AAACAATGTCACCATCCGGTACAGGAACACGGGGCAGAGGGGGTCTCGCAGACTACC	192
132	Db	132	Ala	tyrAspLeuArgValIlysIysThrAspProGlySerLeuGlyLe-	148
193	QY	193	CCGGGTCTCAAAATCCTACTCTGGATATGTCGACACCTCTCCC---	249	
149	Pb	149	ProGlyValIysGlnIyrThrGlyTyrLeuAspAsnGluAsnAspIlyHisLeuPhe	168	
250	QY	250	TTCTGGTTCTCGAAGCCAGACATAACCCAGAAACGTCACCTATCACATTTGGTTGAAT	309	
169	Db	169	TyrTrpPhePheGluSerArgAsnAspProGluAsnAspProValValLeuIyrpLeuAsn	188	
310	QY	310	GGTGGCCCTCGAAGCGATCTTTGATCGGTCTCTTCGAAGAGTGTGGCCCTTCCCATGTC	369	
189	Db	189	GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerIle	208	
370	QY	370	AATTCGACTTTTGATGACTACATCAACCCCTCACTCTGTCGAACGAGGTCCTCCATTACTA	429	
209	Db	209	AsnIlysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValIle	228	
430	QY	430	TTCTGTGCCACGCAATGGGAGTCGGCTTTTCATATAGTCATACGGTTGATGGTCCATT	489	
229	Db	229	PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla-	244	
490	QY	490	AACCCGTGAATGGGGTCGTGCGAAATTCGAGCTTTCAGAGGTTTCAGGGCCGCTACCCA	549	
244	Db	244	-----	244	
550	QY	550	ACCATTGATGCCACTCTCTGATCGATACCTACCAATCTTCCGCAGAGCGCGTTGGAGATC	609	
245	Db	245	-----ValserAspThrValalaAlaGlyLysAspValTyrAlaLeu	258	
610	QY	610	CTGCAAGGATTCCTTGTAGTGACTACCTAGCTTGGACTCTAGGTGTCAGTCTAAGGACTTC	669	
259	Db	259	LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe	274	
670	QY	670	AGTCTATGACGCGAGCTATGGAGGCACTATGGTCTCGATTCTTCAATCAATTTTATC	729	
275	Db	275	HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu	294	
730	QY	730	GAGCAAGATGACAGAAATTCGCCACCGTAGTGTAAATGGTTCTACGCTTAATTCACATCT	789	
295	Db	295	SerHisLysLysArg-----AsnIleAsnLeuGlnSer	305	
790	QY	790	CTGGGAATTTAAACGGCAATCATCGACGAGGCGATTCACGGCCCTTACTACCTCAATTC	849	
306	Db	306	ValLeuIleGlyAsnGlyLeuThrAspGlyTyrThrGlnTyrGluTyrArgProMet	325	
850	QY	850	GCTGTGAACAATACCTACGGTATCAGGCTGTCACAGAGACCGCTACACATACATGAAG	909	
326	Db	326	AlaCysGlyAspGlyGlyTyrProAlaValLeuAspGluSerSerCysGlnSerMetAsp	345	
910	QY	910	TTTGGCAACCAATGCAAAATGGTTGCCAGGATTTGATTTCACCTCGCAACACAGCAACAC	969	

RESULT 7

US-08-967-149-2
Sequence 2, Application US/08967149
Patent No. 5939305
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie Sue
APPLICANT: Thompson, Sheryl Ann
TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF
TITLE OF INVENTION: ASPERGILLUS NIGER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 59393050 No. 5939305disk of No. 5939305th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,149
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,452
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 08/309,341
FILING DATE: 20-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4247.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:


```

Db 346 -----AsnAlaLeuProArg---CysGlnSerMetIleGluSerCysTyrSerSerGlu 362
QY 970 GCGACCGCATAGCTAGCTAGCCCTCTGCGCGGAGCAACCAACATCTGCGAGGACAAAT 1029
Db 363 SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAla 377
QY 1030 GTTGGAGGGCCATAGCTAGCCCTCTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1089
Db 378 LeuLeuAlaProTyrGlnArg---ThrGlyGlnAsnValTyrAspValArgGlyLysCys 396
QY 1090 GATGACCGG-----ACTCGGCAAGTATTACAAACAATTTCTGGCAAG 1134
Db 397 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 416
QY 1135 GACTCTGTGARGAGCTATCGGGTCAACATCAAC---TACACCCAGTCCCAATAATGAC 1191
Db 417 ProGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 436
QY 1192 GTCTACTACGTTTCCAGCAACAGGCGACTTTGTCTGGGCC-----AACTTCATCGAA 1245
Db 437 IleAsnArgAsnPheLeuPheHisGlyAspTrpMetLysProTyrHisArgLeuValPro 456
QY 1246 GACCTCGGAGGAGATCTTGTCTCTCCCGTGGTGTCTCCCTCATCTAT---GGCGAGGCC 1302
Db 457 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 471
QY 1303 GATTACATCTGCACTGTTTCGGCGGTGAGCGGCTTTCCTCGCTCGGAACTACTCCCAA 1362
Db 472 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTrpProGly 491
QY 1363 GCGCGCCAGTTCGAGCGGAGGTCACGCGCCCTGAAAGTC-----AAC 1407
Db 492 GlnAlaGluTyrAlaSerAlaGluLeuGluAspLeuValIleValAspAsnGluHisThr 511
QY 1408 GCGCTCGAGTATGGGAAACTCGCGAGTATGTAATTTCTCTCTCTCTCTCTCTCTCTAG 1467
Db 512 GlyLysLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 531
QY 1468 GCAGGCGTGAAGTCCCATCTACTACGAGCCATCGCTCCCTCGCAATTTTAAACCGGACT 1527
Db 532 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPheAsnAla-gTrp 551
QY 1528 ATCTTCGGT---TGG 1539
Db 552 LeuGlyGlyGluTrp 556

```

RESULT 8

```

US-08-309-341-4
; Sequence 4, Application US/08309341
; Patent No. 5594119
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLIUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSES: No. 55941190 No. 5594119disk of No. 5594119th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/309,341
; FILING DATE: 16-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.

```

```

; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus Niger
; US-08-309-341-4

```

Alignment Scores:

```

Pred. No.: 9,94e-51 Length: 557
Score: 623.00 Matches: 155
Percent Similarity: 48.25% Conservative: 79
Best Local Similarity: 31.96% Mismatches: 175
Query Match: 20.58% Indels: 76
DB: 1 Gaps: 18

```

US-09-712-338-1 (1-1668) x US-08-309-341-4 (1-557)

```

QY 133 GCAACAATCTCACTCCGTTACAGGAACCGGGGAGAGGGCGTCTCGAGACTACC 192
Db 132 AlatyAspLeuArgValLysLysThrAspProSerSerLeuGlyLys-----Asp 148
QY 193 CCGGGTGTCAATCTCTACTCTGATATGTCGACCTCTCCC---GAGTCCCATCCTTC 249
Db 149 ProGlyValLysGlnTyrThrGlyTyrLeuAspAsnGluAsnAspLysHisLeuPhe 168
QY 250 TTCTGTGTTCTTCAAGCCAGACATAACCCAGAACTCCACCTATCATCATGTGTGTAAT 309
Db 169 TyrTrpPhePheGluSerArgAsnAspProGluAsnAspProValValLeuTrpLeuAsn 188
QY 310 GTTGGCCCTTGGAGCGAATCTTTTGATCGGTCTCTCGAAGAGTTGGGCCCTTCCCATGTC 369
Db 189 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerile 208
QY 370 AATTGCACTTTTGATGACTACATCAACCTCTCTCTGGAACAGAGTCTCCAAATTTACTA 429
Db 209 AsnLysLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValile 228
QY 430 TTCTGTCTCCCGCATCTGGAGTTCGGCTTTTCATATAGTATGATACGTTGATGGTCCATT 489
Db 229 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla-----244
QY 490 AACCTGTAACTGGGGTCTGCGAAAATTTCGAGCTTTGAGGAGTTTCAGGGCGCGGTACCCA 549
Db 244 -----244
QY 550 ACCATTGATGCCACTCTGATCGATCTACCAATCTTGCGCGAGAGCGCGCTTGGGAGATC 609
Db 245 -----ValSerAspThrValAlaAlaGlyLysAspValTyrAlaLeu 258
QY 610 CTGCAAGGATTCCTTAGTGGACTACTAGCTGGACTCTAGGTGCTAGTCTTAAGGACTTC 669
Db 259 LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe 274
QY 670 AGTCTATCGAGAGAGCTATGAGGGGCACTATGCTCGTCATTCTTCAATCATTTTATC 729
Db 275 HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu 294
QY 730 GAGCAGATGAGAGAAATGCCAACGGTAGTGTAAATGTGTTCAGCTTAATTTCAACTCT 789
Db 295 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 305
QY 790 CTGGGAATTATTAACGGCATCATCGACGAGCGGATCCAGGCCCTTACTACCTGAATTC 849
Db 306 ValLeuIleGlyAsnGlyLeuThrAspGlyLeuThrGlnTyrGluTyrArgProMet 325

```

SOFTWARE: Patentin Release #1.0, Version #1.25 (ERO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/608,267
FILING DATE: 28-FEB-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/309,341
FILING DATE: 20-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4247,000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 557 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Aspergillus Niger

US-08-608-267-4

Alignment Scores:
Pred. No.: 9,948-51 Length: 557
Score: 623.00 Matches: 155
Percent Similarity: 48.25% Conservative: 79
Best Local Similarity: 31.96% Mismatches: 175
Query Match: 20.58% Indels: 76
DB: 18 Gaps: 1

US-09-712-338-1 (1-1668) x US-08-608-267-4 (1-557)

QY 133 GCAACAAATGTCCACATCCGATACAGGAACCCGGGCGAGAGGCGCTCTCGAGATACC 192
Db 132 AlATyAspueargVallylsysThrAspProSerSerLeuGlyle-----Asp 148
QY 193 CGGGTGTCAAAATCTTACTCTGATGATGTCGACACCTCTCCC---GAGTCCCATACCTTC 249
Db 149 ProGlyVallylsGlnTyrThrGlyTyrLeuAspAsnGluAsnAspHisLeuPhe 166
QY 250 TTCTGTCTTCCAAACCCACATACCCAGAACTGCCTATCATCATTTGTTGAT 309
Db 169 TyrTrpPheGluSerArgAsnAspProGluAsnAspProValLeuLeuTrpLeuAsn 188
QY 310 GTGGGCGCTGAAAGCGATCTTTGATCGGTCTCTTCGAGAGTGGGCCCTTCCCATGTC 369
Db 189 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerIle 208
QY 370 AATTCCGATTTTGATGACTACATCAACCTCCTCATCTGTTGGAACAGGTCTCAATTTACTA 429
Db 209 AsnLysLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValle 228
QY 430 TTCTGTCTCCAGCCATTGGGAGTCCGCTTTTCATATAGTATGATGATCGTGTGGTCCATT 489
Db 229 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla-----244
QY 490 AACCCCTGTAATGGGGTCTGCGAAATTCGAGCTTTTCGAGAGTTCAGGGCGCGTACC 549
Db 244 -----244
QY 550 ACCATTGATGCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 609
Db 245 -----ValSerAspThrValAlaAlaGlyLysAspValTyrAlaLeu 258
QY 610 CTGCAAGGATTCCTTAGTGAGTACTACCTAGCTTGGAGTCTAGGGTGCAGTCTTAAGACTTC 669
Db 259 LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe 274
QY 670 AGTCTATGACGAGAGAGCTATGAGGGCAGCTATGCTCTGCTGCTTCTCAATTTTAC 729

QY 850 GCTGTGAACAATCACTACGGCTCAAGGCTGTCAACGAGACCGCTCTCAACTACATCAAG 909
Db 326 AlaCysGlyAspGly---GlyTyrProAlaValLeuAspGlu---GlySerCysGlnAla 343
QY 910 TTTGCCAACCAATGCCAAATGGTTGCCAGGATTTGATTTCCACTGCCAAACACACAAAC 969
Db 344 MetAspAsnAlaLeuProArg---CysGlnSerMetIleGluSerCysTyrSerSerGlu 362
QY 970 GCGACCGGATAGTACTACGCTCTGCGCCGAGCCCAACCAACATGTCGAGGAGCAAT 1029
Db 363 SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAla 377
QY 1030 GTTGAGGGCCATCACTACGCTTTGCTGCTGTGTGTGTATGATATTCGGCATCCAT 1089
Db 378 LeuLeuAlaProTyrGlnArg---ThrGlyGlnAsnValTyrAspValArgGlyLysCys 396
QY 1090 GATGACCCG-----ACTCGGCCAAGTTATTACAAACAAATTTCTGCAAG 1134
Db 397 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 416
QY 1135 GACTCTGTCATCGAGCTATCGGCTCAACATCAAC---TACACCCAGTCCCAATATGAC 1191
Db 417 ThrGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 436
QY 1192 GTCTACTACGCTTCCAGCAACAGCGGACTTTGTCTGCGCC-----AACTTCATCGAA 1245
Db 437 IleAsnArgAsnPheLeuPheHisGlyAspTyrMetLysProTyrHisArgLeuValPro 456
QY 1246 GACCTCGAGGAGATCCTGCTCTCCCGTGGGTCTCCCTCATCAT---GGCGACGCC 1302
Db 457 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 471
QY 1303 GATTACATCTGCACTGCTGCGCTGCGGCTGCGCTGCGTCTGCGTCCGAACTACTCCCAA 1362
Db 472 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTrpProGly 491
QY 1363 GCGGCCAGTTCGCAAGCGAGGATACACGCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 1407
Db 492 GlnAlaGluTyrAlaSerAlaLysLeuGluAspLeuValValGluAsnGluHisLys 511
QY 1408 GCGCTCGAGTATGGGAACTCGGAGTATGGTATTTCTCTCTCTCTCTCTCTCTCTCTAG 1467
Db 512 GlyLysLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 531
QY 1468 GCGGCGCATGACTCCCATACTACCGCCATCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 1527
Db 532 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnArgTrp 551
QY 1528 ATCTTCGGT---TGG 1539
Db 552 LeuGlyGlyGluTrp 556

RESULT 9
US-08-608-267-4
Sequence 4, Application US/08/608267
Patent No. 5688663
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie Sue
APPLICANT: Thompson, Sheryl Ann
TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: No. 5688663o No. 5688663disk of No. 5688663th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

Db	275	HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAsnGluLeu	294
QY	730	GAGCAGAATGAGAGATTCCCAACGGTAGTGTTAATGGTTCAGCTTAATTTCAACTCT	789
Db	295	SerHisIysIysArg	305
QY	790	CTGGGAATTATTAAACGGCATCTCGACGAGCGATCCAGGCCCTTACTACCTCAATTC	849
Db	306	ValLeuIleGlyAsnGlyLeuThrAspGlyLeuThrGlnTyrGluTyrArgProMet	325
QY	850	GCTGTGAACAATCACTACGGTATCAAGGCTGTCAACGAGACCGTCTACAACCTACATGAAG	909
Db	326	AlaCysGlyAspGly	343
QY	910	TTTGCCAAACCAATGCCAATCGTTGCCAGGATTGTATTCACCTGCAAAACAGACAAC	969
Db	344	MetAspAsnAlaLeuProArg	362
QY	970	CGCACCGCATTAAGCTAGCTACGCCCTCTGCGCGAAGCCACCACATGTCGAGGACAT	1029
Db	363	SerAla	377
QY	1030	GTTGAGGGCCATACACGCTTTGCTGCTGTGTATGATATTCGCGCATCCATAT	1089
Db	378	LeuLeuAlaProTyrGlnArg	396
QY	1090	GATGACCCG	1134
Db	397	GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys	416
QY	1135	GACTCTGTCAATGAGCGTATCGCGGTCAACATCAAC	1191
Db	417	ThrGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp	436
QY	1192	GTCTACTAGCTTTCCAGCAACAGCGCACTTTGCTGGGCC	1245
Db	437	IleAsnArgAsnPheLeuPheHisGlyAspTrpMetLysProTyrHisArgLeuValPro	456
QY	1246	GACCTCGAGGAGATCCTTGCTCTCCCGTCGGGTCTCCCTCATCTAT	1302
Db	457	GlyLeuLeuGlu	471
QY	1303	GATTACATCTGCAACTGGTTCGGCGGTGAGCGGTTCCCTCGCTCGGAATACTCCCAA	1362
Db	472	AspPheIleCysAsnTrpLeuGlyAsnIysAlaTrpThrGluAlaLeuGluTrpProGly	491
QY	1363	CGCGCCAGTTCCGACGCGAGGCTACAGCCCTCGAAAGTC	1407
Db	492	GlnAlaGluTyrAlaSerAlaLysLeuGluAspLeuValValGluAsnGluHisLys	511
QY	1408	GGCGTCGAGPATCGGGAATCCGAGTAGTGGTAATTTCTCCTCACTCGCTCTATGAG	1467
Db	512	GlyIysIysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly	531
QY	1468	GCAGGCATGAAGTCCCACTACACGAGCCATCGCTCCCTGCAATGTTTAACGGACT	1527
Db	532	GlyGlyHisMetValProMetAspGlnProGluSerLeuGluPhePheAsnArgTrp	551
QY	1528	ATCTTCGGT	1539
Db	552	LeuGlyGlyGluTrp	556

```

RESULT 10
US-08-608-452-4
; Sequence 4, Application US/08608452
; Patent No. 5693510
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
;

```

ADDRESSEE: No. 56935100 No. 5693510disk of No. 5693510th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/608,452
FILING DATE: 28-FEB-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/309,341
FILING DATE: 20-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4247,000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 557 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Aspergillus Niger
US-08-608-452-4

Alignment Scores:
Pred. No.: 9.94e-51 Length: 557
Score: 623.00 Matches: 155
Percent Similarity: 48.25% Conservative: 79
Best Local Similarity: 31.96% Mismatches: 175
Query Match: 20.58% Indels: 76
DB: 1 Gaps: 18

US-09-712-338-1 (1-1568) x US-08-608-452-4 (1-557)

```

Alignment Scores:
Fred. No.:          9.94e-51          Length:          557
Score:             623.00             Matches:         155
Percent Similarity: 48.25%             Conservative:    79
Best Local Similarity: 31.96%           Mismatches:     175
Query Match:       20.58%             Indels:         76
DB:                1                  Gaps:          18

US-09-712-338-1 (1-1668) x US-08-608-452-4 (1-557)

QY      133  GCAAAACATGTCAACATCCGTACAGCAAGAACCCGGGCACAGGCGTCTCGAGACTACC 192
      |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      132  AlATyRAsPieuargVallyLysThyRAsPProSerSerLeuGlyIle-----Asp 148
      |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

QY      193  CCGGGTGTCAAATCTTACTCTGGATATGTGCACACCTCTCCC--GAGTCCCATACCTTC 249
      |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      149  ProGlyVallySglnTyRThrGlyTyRLeuAsPAsnGluAsnAspLysHIsLeuPhe 168
      |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::

QY      250  TTCTCGTCTTTCGAAGCCGACATACCCAGAAACTGCACCTATCATTTGTGGTTCAAT 309
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      169  TyRTrpPheGluSerArgAsnAspProGluAsnAspProValValLeuTrpLeuAsn 188
      |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::

QY      310  GGTGGCCCTTGGAGGCAATTTTGATCGGTCTCTTCGAAGAGTGTGGCCCTTCCCATGTC 369
      |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      189  GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerIle 208
      |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::

QY      370  AATTCGACTTTTGATGATPACATCAACCCCTCACTCTGTGGAACGAGGTCTTCCAATTTACTA 429
      |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      209  AsnLysLysIleGlnProValTyRAsnAspTyRAlaTrpAsnSerAsnAlaSerValIle 228
      |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::

QY      430  TTCTCGTCCAGCCATTGGGAGTCGGCTTTTCATATAGTATGATACGGTTGATGGTCCATT 489
      |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      229  PheLeuAspGlnProValAsnValGlyTyRSerTyRSerAsnSerAla-----244
      |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::

QY      490  AACCCTGTGAAGTGGGGTCTGTGAAAAATTCGAAATTCGAGCTTTGAGGAGTTCAGGCGCGGTA 549
      |||  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      244  -----244

```

RESULT 11
US-08-608-224-4
; Sequence 4, Application US/08608224
; Patent No. 5705376
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57053760 No. 5705376disk of No. 5705376th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,224
; FILING DATE: 28-FEB-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus Niger
; US-08-608-224-4

Alignment Scores:
Pred. No.: 9,94e-51 Length: 557
Score: 623.00 Matches: 155
Percent Similarity: 48.25% Conservative: 79
Best Local Similarity: 31.96% Mismatches: 175
Query Match: 20.58% Indels: 76
DB: 1 Gaps: 18

US-09-712-338-1 (1-1668) x US-08-608-224-4 (1-557)

Qy 133 GCACCAATGTCACCATCCGCTACAGAAACCCGGGCGAGAGGGGCTCTGGCAGACTACC 192
Db 132 AlaTyrAspLeuArgValLysLysThrAspProSerSerLeuGlyIle-----Asp 148
Qy 193 CCGGGTGTCAATCCTACTCTGGATATGTCCGACCTCTCCG---GAGTCCCATACCTTC 249
Db 149 ProGlyValLysGlnTyrThrGlyTyrLeuAspAspAsnGluAsnAspLysHisLeuPhe 168
Qy 250 TTCGGTCTTCGAGCCAGCAGACATACCCAGAACTGCACCTATCACATTCTGGTTCAAT 309
Db 169 TyrTrpPhePheGluSerArgAsnAspProGluAsnAspProValValLeuTrpLeuAsn 188
Qy 310 GTGGCCCTCGAAGCGATTCTTTGATCGGTCTCTTCGAGAGATTGGGCCCTTCGCCATGTC 369
Db 189 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerIle 208
Qy 370 AATTCGACTTTTGTGACTATCATCAACCCCTACTCGTGGACGAGCTCTCCCAATTACTA 429

Db	209	AsnLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValIle	228
QY	430	TTCTCTCCAGCCATTTGGAGTGGCGTCTTTCATATAGTCATACGGTGTGATGGTCCAT	489
Db	229	PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSera	244
QY	490	AACCCCTGTAACTGGGGTCTGCGAAATTCGAGCTTTGCAGGAGTTTCAGGGCGGTACCCA	549
Db	244	-----	244
QY	550	ACCATTTGATGCCACTCTGATCGATACCTACCAATCTTTCGCCGACAGGCCGCTTGGGAGTC	609
Db	245	-----ValSerAspThrValAlaAlaGlyIleAspValTyrAlaLeu	258
QY	610	CTGCAAGGATTCCTTAGTGGACTACCTAGCTTGGACTCTAGGGTCGAGTCTAAGACATTC	669
Db	259	LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe	274
QY	670	AGTCTATGACGAGGAGCTATGGAGGCGACTATGCTCTGCATCTTCTCAATCATTTTAC	729
Db	275	HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluLeuLeu	294
QY	730	GAGCAGAAATGACAGAAATTCGCAACGCTAGTGTATGTTGTTTCAATTAATTCACATCT	789
Db	295	SerHisLysLysArg-----AsnIleAsnLeuGlnSer	305
QY	790	CTGGGAATTAATTAACGGCATCATCGACGAGCGATCCAGGCCCTTACTACCTGAAATC	849
Db	306	ValLeuIleGlyAsnGlyLeuThrAspGlyLeuThrGlnTyrGluTyrTyrArgProMet	325
QY	850	GCTGTGAACAATACCTACGGTATCAAGGCTGTCAACGAGACCGTCTACAACATACATGAAG	909
Db	326	AlaCysGlyAspGly---GlyTyrProAlaValLeuAspGlu---GlySerCysGlnAla	343
QY	910	TTTGCCAAACCAATGCCAAATCGTTGCCAGGATTTGATTTCCACTCGCAACACAGACAAC	969
Db	344	MetAspAsnAlaLeuProArg---CysGlnSerMetIleGluSerCysTyrSerSerGlu	362
QY	970	CGCACCGCATTAGCTGACTACGCCCTCTGCGCGAAGCCCAACATGTCGAGGACAAAT	1029
Db	363	SerAla-----TrpValCysValProIleSerIleTyrCysAsnAla	377
QY	1030	GTTGAGGGCCATACTACGCTTTGCTGCTGCTGTGTATGATATTCGCGATCCATAT	1089
Db	378	LeuLeuAlaProTyrGlnArg---ThrGlyGlnAsnValTyrAspValArgGlyLysCys	396
QY	1090	GATGACCCG-----ACTCCGCCAAGTTATTACAACAATTTCTGGCAAG	1134
Db	397	GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys	416
QY	1135	GACTCTGTCATGACGCTATCGGCGTCAACATCAAC--TACACCCAGTCCCAATGAC	1191
Db	417	ThrGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp	436
QY	1192	GTCTACTACGCTTTCAGCAACAGGCGACTTGTCTGGCCC-----AACTCATCGAA	1245
Db	437	IleAsnArgAsnPheLeuPheHisGlyAspTrpMetLysProTyrHisArgLeuValPro	456
QY	1246	GACCTCGAGGAGATCCTTGTCTCTCCCGTGGCTCTCCCTCATCTAT---GGCGAGCC	1302
Db	457	GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla	471
QY	1303	GATTACATCTGCAACTGGTTTCGGCGGTACAGCGCGTTTCCCTCGCTCGGAATACTCCCA	1362
Db	472	AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTrpProGly	491
QY	1363	CGCGCCAGTTCCGAACGCGAGGTACACGCCCTGAAAGTC-----AAC	1407
Db	492	GlnAlaGluTyrAlaSerAlaLysLeuGluAspLeuValValGluAsnGluHisLys	511
QY	1408	GGCGTCGAGTATGGGAAATCCGAGTATGGTAATTTCTCCTTCACTCGCGTCTAGAG	1467

Db 512 GlyLysLysIleGlyGlnValIysSerHisGlyAsnPheThrPheMetArgLeuTrpGly 531
QY 1468 GRAGCCATGAAGTCCCATACTACCAAGCCATCGCCTCCCTGCAATTGTTTAAACGGACT 1527
Db 532 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnArgTrp 551
QY 1528 ATCTTCGGT---TGG 1539
Db 552 LeuGlyGlyGluTrp 556

RESULT 12

RESULT 12
 US-08-967-149-4
 ; Sequence 4, Application US/08967149
 ; Patent No. 5939305
 ; GENERAL INFORMATION:
 ; APPLICANT: Yaver, Debbie Sue
 ; APPLICANT: Thompson, Sheryl Ann
 ; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF
 ; TITLE OF INVENTION: ASPERGILLUS NIGER
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 5939305o No. 5939305disk of No. 5939305th America, Inc.
 ; STREET: 405 Lexington Avenue, Suite 6400
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/967,149
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/608,452
 ; FILING DATE: 28-FEB-1996
 ; APPLICATION NUMBER: US 08/309,341
 ; FILING DATE: 20-SEP-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney, Karen A.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 4247-000-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 867 0123
 ; TELEFAX: 212 867 0298
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 557 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; ORIGINAL SOURCE:
 ; ORGANISM: Aspergillus Niger
 US-08-967-149-4

Alignment Scores:

Pred. No.:	9.94e-51	Length:	557
Score:	623.00	Matches:	155
Percent Similarity:	48.24%	Conservative:	795
Best Local Similarity:	31.94%	Mismatches:	175
Query Match:	20.58%	Indels:	76
DB:	2	Gaps:	18

US-09-712-338-1 (1-1668) x US-08-967-149-4 (1-557)

[illegible]

[illegible]

193 CCGGGTGTCAAACTCTACTCTGGATATGTCGACAGCTCTCC---GAGTCCCATACCTTC 249
149 ProGlyValLysGlnTyrThrGlyTyrLeuAspAsnGluAsnAspLysHisLeuPhe 168
250 TTCTGGTCTTCGAGCAGACATACCCAGAACTGCACCTATCACATCTATCACATCTGGTGTGAAT 309
169 TyrTrpPheGluSerArgAsnAspProGluAsnAspProValValLeuTrpLeuAsn 188
310 GGTGGCCCTCGAAGCGATTCTTTGATCGGTCTCTTCGAAGAGTTGGGCCCTTGGCCATGTC 369
189 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerIle 208
370 AATTTCGACTTTTGTGATGACTACATCAACCTCTACTCGTGGAAACGAGGTCTCCCAATTACTTA 429
209 AsnLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValIle 228
430 TTCCTGTCGCCAGCCATTGGGAGTCGGCTTTTCATATAGTAGATGATGAGGTTCATGGTCCATT 489
229 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla----- 244
490 AACCTGTAACTGGGGTGGTCGAAAATTTCGAGCTTTGCAGGAGTTGCAGGCCGGTACCCA 549
244 ----- 244
550 ACCATTGATGCACTCTCTGATCGATACACTACCAATCTTCGCCGAGAGCCGCTTGGAGATC 609
245 -----ValSerAspThrValAlaAlaGlyLysAspValTyrAlaLeu 258
610 CTGCAAGGATTCCTTAGTGCATCTACCTAGCTTGGAGTCTAGGCTGAGTCTAAGGAGCTTC 669
259 LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe 274
670 AGTCTATGACGAGCAGAGCTATGAGGGCAGCATGGTCCTGSCATCTTCAATCATTTTATC 729
275 HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu 294
730 GAGCGAATGAGAGAATTGCCAGGTAGTGTAAATGTTGTTAGTGTTCAGCTTAATTCACACTCT 789
295 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 305
790 CTGGGAATTTATACCGCATCATCGAGCGCATCCAGGCCCTTACTACCTGAAATTC 849
306 ValLeuIleGlyAsnGlyLeuThrAspGlyLeuThrGlnTyrGluTyrTyrArgProMet 325
850 GCTGTGAACAATACCTACGCTATCAAGCTGTCAACGAGACCGTCTCAACTACATCAAG 909
326 AlaCysGlyAspGly---GlyTyrProAlaValLeuAspGlu---GlySerCysGlnAla 343
910 TTTGCCAACCAATGCGCAATGGTTGCCAGGATTTGATTTCACCTGCAACACAGCAAC 969
344 MetAspAsnAlaLeuProArg---CysGlnSerMetIleGluSerCysTyrSerSerGlu 362
970 CGCACCGCATAGCTGACTACGCCCTCTGGCGCGAAGCCACCAATCTGAGGAGACAAT 1029
363 SerAla-----TyrPheCysValProAlaSerIleTyrCysAsnAsnAla 377
1030 GTTGGAGGGCCATCTACTACGCCCTTGGCTGGTGTGTATGATATTCCGATCCATAT 1089
378 LeuLeuAlaProTyrGlnArg---ThrGlyGlnAsnValTyrAspValArgGlyLysCys 396
1090 GATGACCCG-----ACTCCGCCAGTTATTACACAAATTTCTGCGAAG 1134
397 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 416
1135 GACTCTGTCATGGAGCTCTGGCGGTCAACATCAAC---TACACCCAGTCCCAATATGAC 1191
417 ThrGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 436
1192 GTCTACTACGCTTTCCAGCAACAGCGGACTTTGTCTGGCCC-----AACTTCATCGAA 1245
437 IleAsnArgAsnPheLeuPheHisGlyAspTrpMetLysProTyrHisArgLeuValPro 456
-----GGCGACGCC 1302

```

Pred. No.: 1,486-45 Length: 491
Score: 568.50 Matches: 151
Percent Similarity: 44.65% Conserved: 66
Best Local Similarity: 31.07% Mismatches: 184
Query Match: 18.78% Indels: 85
DB: 1 Gaps: 14

US-09-712-338-1 (1-1668) x US-09-640-305-4 (1-491)

QY 133 GCAACAATGTCACCATCCGGTACAAAGAACCCGGGCGAGAGCGGCTCGGAGACTACC 192
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
62 AlaTyrSerLeuArgIleLysProLeuAspProLysSerLeuGlyVal----- 77
QY 193 CCGGCTGCAAACTCTGATCTGGAATGTCGACACCTCTCCCGAGTCCCATACCTTCTTC 252
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
78 AspThrValLysGlnTrpSerGlyTyrLeuAspTyrGlnAspSerLysHisPhePheTyr 97
QY 253 TGGTCTTTCGAAGCCACACATAACCCAGAACTCCACCTATCACATTGCTGTTCAATGGT 312
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 TrpPheGluSerArgAsnAspProGluAsnAspProValIleLeuTrpLeuAsnGly 117
QY 313 GGCCTGGAAGCGATTCTTTGATCGGTCTCTTCGAGAGTTGGGCCCTTCCCATGTCAAT 372
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
118 GlyProGlyCysSerPheValGlyLeuPheGluLeuGlyProSerSerIleGly 137
QY 373 TCAGATTTCGATGACTACATCAACCTCTACCTGCGAACGAGGTCTCCAAATTTACTATTC 432
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
138 AlaAspLeuLysProIleTyrAsnProTyrSerTrpAsnSerAsnAlaSerValIlePhe 157
QY 433 CTCTCCAGCATTTGGGAGTCGCTTTTCATATAGTATGATACGGTTGATGGGTCATTAAC 492
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
158 LeuAspGlnProValGlyValGlyPheSerTyrGlyAsp----- 170
QY 493 CTTGTAATCGGGTCTGCGAAAATTCGAGCTTTCGAGGAGTTCAGGCGCGGTACCCCAAC 552
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
170 ----- 170
QY 553 ATGTATGCACTCTGATCGATACCAATCTTTCGCGAGAGCGCGCTTGGGAGATCCTG 612
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
171 -----SerLysValSerThrAspAspAlaAlaLysAspValTyrIlePheLeu 187
QY 613 CAAGGATTCCTTAGTGGACTACTAGCTTGGACTCTAGGTGCGAGTCAAGGACTTCAGT 672
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
188 AspLeuPheGluArgPheProHisLeu-----ArgAsnAspPheHis 203
QY 673 CTATGACGAGAGCATGATGAGGCGACTATGCTGCTGATCTTCAATCATTTTACGAG 732
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
204 IleSerGlyGluSerTyrAlaGlyHisTyrLeuProLysIleAlaHis----- 219
QY 733 CAGATGAGAGAAATGCCAAGGTAGTGTAAATGGTTCAGCTTAATTTCAACTCTCTG 792
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
220 -----GluIleAlaValHisAlaGluAspSerPheAsnLeuSerSerVal 236
QY 793 GGAATTTATTAACGGCATCTCGACGAGGCGATCCAGGCCCTTACTACCTGAATTCGT 852
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
237 LeuIleGlyAsnGlyPheThrAspProLeuThrGlnTyrGlnTyrGluProMetAla 256
QY 853 GTGAACAATPACCTACGGTATCAAGGCTGTCAACGAGACCGTCTACAACTACATGAATTT 912
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
257 CysGlyGluGly---GlyTyrProAlaValLeuGlu----- 267
QY 913 GCCAACCAATGCCAAATGGTTCAGGAT----- 942
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
268 -----ProGluAspCysLeuAspMetAsnArgAsnLeuProLeuCysLeuSer 283
QY 943 TTGATTTCCACTCGCAACAGACAAACCCAGCATTAGCTAGTACGCCCTCTCGGCC 1002
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
284 LeuValAspArgCysTyrLysSerHis-----SerValPheSerCysVal 298
QY 1003 GAAGCCCAACATGTGAGGAGCAATTTGAGGGCCATATACGCTTTCCTGCTGCT 1062
Dy ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 LeuAlaAspArgTyrCysGluGlnIleThrGly---ValTyrGluLysSerGlyArg 317

```

RESULT 14

```

US-08-360-673-4
; Sequence 4, Application US/08360673
; Patent No. 5679544
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360.673
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US

```

TELECOMMUNICATION INFORMATION:

TELEPHONE: (610)454-3839
TELEFAX: (610)454-3808
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-360-673-4

Alignment Scores:
Pred. No.: 1,48e-45 Length: 491
Score: 568.50 Matches: 151
Percent Similarity: 44.65% Conservative: 66
Best Local Similarity: 31.07% Mismatches: 184
Query Match: 18.78% Indels: 85
DB: 1 Gaps: 14

US-09-712-338-1 (1-1668) x US-08-360-673-4 (1-491)

QY	133	GCAACAATGTCACCATCGGTACAGGAGCCGGGAGAGGGGCTCTGCGAGACTACC	192
DB	62	AlafyrSerLeuArgilelySerProLeuAspProLysSerLeuGlyVal	77
QY	193	CGGGGTGCAAACTCTACTCTGATATGTCGACACCTCTCCGAGTCCCATACCTTC	252
DB	78	AspThrValylsGlnTrpSerGlyTyrLeuAspTyrGlnAspSerlyshisPhePheTyr	97
QY	253	TGGTCTTCGAGCCAGACATACCCAGAACTGCACCTATCACATGTGGTGTGATG	312
DB	98	TrpPhePheGluSerArgAsnAspProGluAsnAspProValileLeuTrpLeuAsnGly	117
QY	313	GGCCCTGGAGCCATCTTGTGCTCTCTCGAGAGTGGCCCTCTGCCATGTCAT	372
DB	118	GlyProGlyCysSerPheValGlyLeuPhePheGluLeuGlyProSerSerleGly	137
QY	373	TCGACTTTTGATGACTACATCAACCTCTCTGGAAGAGGTCTCCAAATTTACATTC	432
DB	138	AlaAspLeuLysProIleTyrAsnProTyrSerTipAsnSerAsnAlaSerValilePhe	157
QY	433	CTGTCGCCAGCANTGGAGTGGCTTTTATATAGTACAGCTTATGGTCCATTAAC	492
DB	158	LeuAspGlnProValGlyValGlyPheSerTyrGlyAsp	170
QY	493	CTGTGTAACCTGGGTCTCGAABATTCGAGCTTTGCGAGGAGTTCAGGCGCGGTACCCAAAC	552
DB	170		170
QY	553	ATTGATGCCACTCTGATCGATACCAATCTTTCGCGAGAGCCGGCTTGGGAGATCCTG	612
DB	171	-----SerlySValSerThrAspAlaAlaLysAspValTyrIlePheLeu	187
QY	613	CAGGATTCCTTAGTGACTACCTAGCTTGGACTCTAGGCTGAGCTTAAGGACTTCAGT	672
DB	188	AspLeuPhePheGluArgPheProHisLeu-----ArgAsnAsnAspPheHis	203
QY	673	CTATGGAGGAGACTATGAGGCGACTATGCTGCTGCTCATTCATCATCTTTTACGAG	732
DB	204	IleSerGlyGluSerTyrAlaGlyHisTyrLeuProLysIleAlaHis-----	219
QY	733	CAGAAATGAGAAATGCCAGCGTAGTGTAAATGGTGTTCAGCTTAATTTCAACTCTCTG	792
DB	220	-----GluIleAlaValValHisAlaGluAspSerSerPheAsnLeuSerSerVal	236
QY	793	GGAATTATTACGGCATCATCGAGCGCATCCAGGCCCTTACTACCTCCCTGATTCGCT	852
DB	237	LeuileGlyAsnGlyPheThrAspProLeuThrGlnTyrGlnTyrGluProMetAla	256
QY	853	GTGAACAATACCTAGGCTATCAAGGCTCTCAACGAGACCGTCTCAACTACATGAGTTT	912
DB	257	CysGlyGluGly---GlyTyrProAlaValLeuGlu-----	267

QY	913	GCCAACCAATGCCAAATGTTGCGAGAT-----	942
DB	268	-----ProGluAspCysLeuAspMetAsnArgAsnLeuProLeuCysLeuSer	283
QY	943	TTGATTTCCACCTGCAAAACAGACAAACCGCACCATTTAGCTAGCTAGCCCTCTGCGCC	1002
DB	284	LeuValAspArgCysTyrLysSerHis-----SerValPheSerCysVal	298
QY	1003	GAAGCCACCAACATGTGCGAGGAGCAATGTTGAGGGGCCCATCTAGCCCTTTGCTGCTGT	1062
DB	299	LeuAlaAspArgTyrCysGluGlnGlnIleThrGly---ValTyrGluLysSerGlyArg	317
QY	1063	GGTGTGTATGATATTCGG-----CATCATATGATGATGACCCG-----ACT	1101
DB	318	AsnProTyrAspIleArgSerLysCysGluAlaGluAspAspSerGlyAlaCysTyrGln	337
QY	1102	CGCCCAAGTTATTACACAAATTTCTGGCAAAAGACTCTGTCATCGAGCTATCGCGCTC	1161
DB	338	GluGluIleTyrIleSerAspTyrLeuAsnGlnGluValGlnArgAlaLeuGlyThr	357
QY	1162	AACATCAACTACCCAGTCC---AATAATGACGCTCTACTAGCTTTCAGCAAAACAGCC	1218
DB	358	AspValSerSerPheGlnGlyCysSerSerAspValGlyIleGlyPheAlaPheThrGly	377
QY	1219	GACTTGTGTCGCCCAACTTCATCGAAGACCTCGAGGAGATCCTTGTCTCTCCCGTGC	1278
DB	378	AspGlyProSerPro---PheHisGlnTyrValAlaGluLeuAspGlnAspIleAsn	396
QY	1279	GTCCTCCCTCATCTATGGCGACCCGCTTACATCTGCACTGCTGGTGGCGGTGAGGCGGT	1338
DB	397	ValLeuIleTyrAlaGlyAspLysAspTyrIleCysAsnTrpLeuGlyAsnLeuAlaTrp	416
QY	1339	TCCCTCGCTGCGAACTACTCCCAAGCCGCCAGTTCGAGCGCGGTACAGCCCGCTG	1398
DB	417	ThrGluLysLeuGluTrpArgTyrAsnGluGluTyrLysLysGlnValLeuArgThrTrp	436
QY	1399	AAAGTCAACGGCGTCTGAG-----TATGGGAAATCTCGAGATGATGATGATTTCTCCTTC	1452
DB	437	LysSerGluGluThrAspGluThrIleGlyGluThrLysSerTyrGlyProLeuThrTyr	456
QY	1453	ACTCGCGTCTATGAGCGAGCCCATGAAAGTCCCATATACAGCCCATTCGCTCCCTGCAA	1512
DB	457	LeuArgIleTyrAspAlaGlyHisMetValProHisAspGlnProGluAsnSerLeuGln	476
QY	1513	TGCTTTAACGGGACTATC 1530	
DB	477	MetValAsnSerTrpIle 482	

RESULT 15

US-08-999-324-33
Sequence 33, Application US/08999324
Patent No. 5945329
GENERAL INFORMATION:
APPLICANT: Bredham, Klaus
APPLICANT: Keiland-Brandt, Morten
APPLICANT: Mortensen, Uffe
APPLICANT: Olesen, Kjeld
APPLICANT: Stennicke, Henning
APPLICANT: Wagner, Fred
TITLE OF INVENTION: CUSTOMIZED PROTEASES
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
STREET: 3100 No. 5945329west Center, 90 S. 7th Street
CITY: Minneapolis
STATE: MN
COUNTRY: U.S.A.
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,324
FILING DATE: 23-JUL-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/329,892
FILING DATE: 27-OCT-1994
APPLICATION NUMBER: 08/144,704
FILING DATE: 28-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kettleberger, Denise M
REGISTRATION NUMBER: 33,924
REFERENCE/DOCKET NUMBER: 8648.44UCS1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/332-5300
TELEFAX: 612/332-9081
TELEX:

```

1  / RELEX:
2  / INFORMATION FOR SEQ ID NO: 33:
3  /
4  / SEQUENCE CHARACTERISTICS:
5  /     LENGTH: 532 amino acids
6  /     TYPE: amino acid
7  /     STRANDEDNESS: single
8  /     TOPOLOGY: linear
9  /
10 /     MOLECULE TYPE: protein
11 /     HYPOTHEICAL: NO
12 /
13 /     ANTI-SENSE: NO
14 /     FRAGMENT TYPE: internal
15 /     ORIGINAL SOURCE:
16 /     US-08-899-324-33

```

Alignment Scores:	
Pred. No.:	3.08e-41
Score:	523.50
Percent Similarity:	43.60%
Best local Similarity:	23.05%
Query Match:	17.25%
DB:	2
Length:	532
Matches:	147
Conservative:	74
Mismatches:	200
Indels:	85
Gaps:	19

US-09-712-338-1 (1-1668) x US-08-899-324-33 (1-532)

QY	97	CCCAGAAACCCACC	CGGGGTCAAGACT	-----CCTACAACCGCAACAAT	141
			:::		
Db	82	ProLysPheProGluAlaIleLysThrLysAspPheValVallYsAsnAsp	101		
QY	142	GTCACCATCCGGTAC	-----AAGGAACCCGGGCAGAGGCCGC	180	
Db	102	AlaIleGluAsnTyrlinLeuArgValAsnLysIleLysAspProLysIleLeuGlyIle	121		
QY	181	TGCAGAGACTACCCGGGTGTCAAATCCTACTCTCGATATGTGCAACCTCTCCCGAG---	237		
Db	122	-----aspProAsnValThrGlnTyrrThrGlyrThrGlyrLeuAspValGluAspGluAsp	138		
QY	238	TCCCATACCTTTCTTGTTCTTCGAGCCAGACATAACCCAGAAGAACTGCACCTATCAC	297		
Db	139	LysHisPhePheThrTrpPheGluSerArgAsnAspProAlaLysAspProValIle	158		
QY	298	TTGTGGTTGAATGGTGCCCTGGAGCGAANTCTTTGATCGGTCTCTCGAAGAGTTGGC	357		
Db	159	LeutrPrLeuAsnGlyGlyProGlyCysSerSerLeuthrGlyLeupheGluLeuGly	178		
QY	358	CCTTGCCATGTCAAATTGCAGCTTTTGATGACTACATCAACCTCACTCGTGGAACAGAGTC	417		
Db	179	ProSerSerIleGlyProAspLeuLysProIleGlyAsnProTyrrSerTipAsnSerAsn	198		
QY	418	TCCAATTACTATTCTGTGCCACCAATGGGAGTCGGCTTTTCATATAGTAGTACGGTT	477		
		:::			
Db	199	AlaThrValIlePheLeuAspGlnProValasnValGlyPheSerTyrrSer-	215		
QY	478	GATGGGTCCATTAAACCTGTGAATCGGTGGGTGCTCGAAAATTCGAGCTTTGCAGGAGATT	537		
Db	216	-----GlyserSer	218		
QY	538	GGCGGTACCCACCAATTGATGCCACTCTGATCGATACTACCAATCTTGGCGCAGAGCC	597		

[illegible]

Search completed: November 21, 2003. 17.46.59

Mon Nov 24 13:41:03 2003

us-09-712-338-1.rai

Page 18

Job time : 52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 21, 2003, 14:49:17 ; Search time 62 Seconds
(without alignments)
8540.514 Million cell updates/sec

Title: US-09-712-338-1

Perfect score: 3027

Sequence: 1 atcggtggtacgaattctt.....ccagtggtggtatgcatag 1668

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DBV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09712338/runat 17112003.170148.10497/app.query.fasta_1.1863
-DB=A Geneseq 19Jun03 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @C@N 1 1 92 @runat 17112003.170148.10497 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A Geneseq 19Jun03: **
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: **
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: **
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: **
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: **
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: **
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: **
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: **
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: **
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: **
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: **
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: **
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: **
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: **
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: **
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: **
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: **
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: **
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: **
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: **
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: **
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: **
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: **
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: **
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: **

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2975	98.3	554	19	AAW56099	A. oryzae ATCC2038
2	2376	78.5	551	23	ABR38864	A. niger serine ca
3	980.5	32.4	623	23	ABR38865	A. niger serine ca
4	947	31.3	554	23	ABR38859	A. niger serine ca
5	925	30.6	526	23	ABR38817	A. niger serine ca
6	636	21.0	508	23	ABG32049	S. cerevisiae BAX-
7	630	20.8	536	23	ABR38819	A. niger carboxype
8	620	20.5	556	17	AA96737	A. niger Bo-1 carb
9	568.5	18.8	491	15	AA96738	A. niger SPAG 2 ca
10	541.5	17.9	550	23	ABG93281	Sequence of protea
11	419	13.8	481	23	ABR38845	C. albicans BAX-as
12	383	12.7	482	21	ABG30065	A. niger carboxype
13	372	12.3	502	21	ABR38843	Arabidopsis thalia
14	362.5	12.0	476	22	AA988381	Arabidopsis thalia
15	361.5	11.9	479	21	AAW72966	A. niger carboxype
16	361.5	11.9	479	21	AAW72966	Human membrane or
17	360.5	11.9	486	21	AAW72966	Human serine carbo
18	360.5	11.9	486	21	AAW72966	Arabidopsis thalia
19	360.5	11.9	486	21	AAW72966	Arabidopsis thalia
20	360.5	11.9	486	21	AAW72966	Amino acid sequenc
21	360.5	11.9	476	21	AAW72966	Human TANGO 176.
22	360.5	11.9	476	22	AAU29228	Human PRO polypept
23	360.5	11.9	476	22	AAU29228	Human hydrophobic
24	360.5	11.9	476	22	AAU29228	Human PRO223 prote
25	360.5	11.9	476	22	AAU29228	Novel human secret
26	360.5	11.9	476	24	ABU71316	Human PRO223 prote
27	360.5	11.9	476	24	ABU71316	Human PRO polypept
28	360.5	11.9	476	24	ABU71316	Human secreted/tra
29	360.5	11.9	476	24	ABU71316	Human secreted/tra
30	360.5	11.9	476	24	ABU71316	Novel human secret
31	360.5	11.9	476	24	ABU71316	Human secreted/tra
32	360.5	11.9	476	24	ABU71316	Human secreted/tra
33	360.5	11.9	476	24	ABU71316	Human secreted/tra
34	360.5	11.9	476	24	ABU71316	Human secreted/tra
35	360.5	11.9	476	24	ABU71316	Human PRO polypept
36	360.5	11.9	476	24	ABU71316	Human PRO polypept
37	360.5	11.9	476	24	ABU71316	Human secreted/tra
38	360.5	11.9	476	24	ABU71316	Human secreted/tra
39	360.5	11.9	476	24	ABU71316	Human secreted/tra
40	360.5	11.9	476	24	ABU71316	Human secreted/tra
41	360.5	11.9	476	24	ABU71316	Human secreted/tra
42	360.5	11.9	476	24	ABU71316	Human secreted/tra
43	360.5	11.9	476	24	ABU71316	Human secreted/tra
44	360.5	11.9	476	24	ABU71316	Human secreted/tra
45	360.5	11.9	476	24	ABU71316	Human secreted/tra

ALIGNMENTS

RESULT 1
AAW56099
ID AAW56099 standard; Protein; 554 AA.
XX
AC AAW56099;
XX
DT 27-AUG-1998 (first entry)
XX
DE A. oryzae ATCC20386 carboxypeptidase I protein.
KW Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous;
KW food industry.
XX
OS Aspergillus oryzae.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT /label= signal

	Protein	19..554 /label= carboxypeptidase I			
FT	XX	WO9814599-A1.			
FF	XX				
PD	XX	09-APR-1998.			
PX	XX	03-OCT-1997;	97WO-US17977.		
PR	XX	27-NOV-1996;	96US-0757534.		
PS	XX	04-OCT-1996;	96US-0726880.		
PT	XX	(NOVO) NOVO NORDISK BIOTECH INC. (NOVO) NOVO-NORDISK AS.			
PI	XX	Berka R, Blinkovsky A, Brown K, Dambmann C, Golightly E; Klotz A, Mathisen TE, Rey M;			
DR	XX	WPI; 1998-240098/21. N-PSDB; AAU28620.			
CC	XX	Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous substrates, useful for improving flavour of foods			
CS	XX	Claim 1; Fig 3; 82pp; English.			
CY	XX	This sequence represents carboxypeptidase I from Aspergillus oryzae. This polypeptide has an optimal activity in the range of pH 3.0-7.5 at 25 deg. C, optimal activity in the range of 55-60 deg. C at pH 4. and a residual activity of at least 65.5% after 30 minutes at pH 4.0 and 60 deg. C. It also has the capacity to hydrolyse X from N-CBZ-Ala-X where N-CBZ is N-carbobenzoxy and X is any amino acid. The carboxypeptidases can be used for obtaining hydrolisates (which can be enriched in free glutamic acid or peptide bound glutamic acid residues) from proteinaceous substrates. The carboxypeptidases can be used in flavour-improving compositions in the food industry. The products can also be used for the production of polypeptides free of carboxypeptidase activity.			
DQ	XX	SQ Sequence 554 AA;			
EI	XX	Alignment Scores:			
FJ	XX	Pred. No.: 1.09e-259 Length: 554			
GK	XX	Score: 2975.00 Matches: 554			
HG	XX	Percent Similarity: 100.00% Conservative: 0			
IH	XX	Best Local Similarity: 100.00% Mismatches: 0			
JL	XX	Query Match: 98.28% Indels: 0			
KM	XX	DB: 19 Gaps: 0			
LN	XX	US-09-712-338-1 (1-1668) x AAWS6099 (1-554)			
MN	QY	1 ATGGGTGGTCAAGAAATTCCTCAGTGACTACCTTTGGTGGCGGCCCTTCCA 60			
NN	Db	1 MetArgGlyTyrgluPheLeuSerValLeuproLeuAlaAalaSerTrpAlaleuPro 20			
OQ	QY	61 GGAAATAACCGGCCTCGGTGAGTAGAGAGCTACCACAAGAACCACCGGGGTCAAG 120			
OR	Db	21 GlySerThrProalaserValglyargarginleuprolysasnProthrGlyValys 40			
OS	QY	121 ACTCTTAACACCGCAAATGTCACCATCCGGTACAGGAACCCGGGGGAGAGGGGTC 180			
OT	Db	41 ThrLeuthrThralaalasnasnvalthrileargtyrlsysgluproglyalagluglyval 60			
OU	QY	161 TCCGAGACTACCCGGGTGTCAANCTACTCTGGATATGTCACCACTCTCCCAGATCC 240			
OV	Db	61 CysGlutThrProglyValylserTyrSerGlyTyrValaspThrSerProgluser 80			
OW	QY	241 CATACCTCTCTGGTTCATTGGAAGCCAGACATAACCAGAAAACCTCACCTATCACATTG 300			
OX	Db	81 HisThrPhePhetrpPheGluaArghisasnProgluthralaproillethrLeu 100			
OY	QY	301 TG GTTGAATGGTGGCCCTCGAAGCGATTCCTTCATCGCTCTCTTCGAGAGGTTGGCCCT 360			
PA	Db	101 TrpLeuasnglyvlglyProGlySerAspSerLeulleGlyLeuphegluglueudglyPro 120			

QY 1441 AATTCTCTCTCACTCGGCTATAGGACGGCCATGAAGTCCCATACTACCGCCCATC 1500
 Db |||||
 QY 481 AsnPheSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProile 500
 |||||
 QY 1501 GCCTCCCTGCAATTGTTAAACCGGACTATCTCGGTTGGGATATCGCAGAGGGCCAGAAG 1560
 Db |||||
 QY 501 AlaserLeuGlnLeuPheAsnArgThrIlePheGlyTrpAspIleAlaGluGlyGlnLys 520
 |||||
 QY 1561 AAGATCTGGCCCACTACAAGACGAATGGAACGGCTACAGCTACGCATACACATCGCTGCC 1620
 Db |||||
 QY 521 LysIleTrpProSerTyrLysThrAsnGlyThrAlaThrAlaThrHisThrGlnSerSer 540
 |||||
 QY 1621 GTCCGGCTGCTACGGCTACAGCATGTCCAGTCTTGGTATG 1662
 Db |||||
 QY 541 ValProLeuProThrAlaThrSerMetSerSerValGlyMet 554
 |||||

RESULT 2
 ABR38864
 ID ABR38864 standard; Protein; 551 AA.
 AC ABR38864;
 XX
 XX DT 24-APR-2003 (first entry)
 XX
 XX DE A. niger serine carboxypeptidase polypeptide #3.
 XX
 XX KW Protease; fungal infection; aspergilliosis; food; tanning; detergent;
 KW protein solubility; viscosity; taste; texture; nutritional value;
 KW EC3.4.16.6.
 XX
 XX OS Aspergillus niger.
 XX
 XX PN WO200268623-A2.
 XX
 XX PD 06-SEP-2002.
 XX
 XX PF 22-FEB-2002; 2002WO-EP01984.
 XX
 XX PR 23-FEB-2001; 2001EP-0200657.
 PR 23-FEB-2001; 2001EP-0200658.
 PR 23-FEB-2001; 2001EP-0200660.
 PR 26-FEB-2001; 2001EP-0200706.
 PR 26-FEB-2001; 2001EP-0200707.
 PR 26-FEB-2001; 2001EP-0200708.
 PR 26-FEB-2001; 2001EP-0200719.
 PR 28-MAR-2001; 2001EP-0000075.
 PR 28-MAR-2001; 2001EP-0000078.
 PR 28-MAR-2001; 2001EP-0000080.
 PR 28-MAR-2001; 2001EP-0000087.
 PR 28-MAR-2001; 2001EP-0000088.
 PR 21-MAY-2001; 2001EP-0000156.
 PR 21-MAY-2001; 2001EP-0000159.
 PR 21-MAY-2001; 2001EP-0000160.
 PR 21-MAY-2001; 2001EP-0000162.
 PR 21-MAY-2001; 2001EP-0000165.
 PR 21-MAY-2001; 2001EP-0000166.
 PR 21-JUN-2001; 2001EP-0000168.
 PR 21-JUN-2001; 2001EP-0000240.
 PR 21-JUN-2001; 2001EP-0000242.
 PR 21-JUN-2001; 2001EP-0000244.
 PR 21-JUN-2001; 2001EP-0000246.
 PR 12-JUL-2001; 2001EP-0000280.
 PR 12-JUL-2001; 2001EP-0000285.
 PR 30-JUL-2001; 2001EP-0000323.
 PR 30-JUL-2001; 2001EP-0000327.
 PR 02-AUG-2001; 2001EP-0000341.
 PR 02-AUG-2001; 2001EP-0000342.
 PR 02-AUG-2001; 2001EP-0000343.
 PR 02-AUG-2001; 2001EP-0000344.
 PR 09-AUG-2001; 2001EP-0000357.
 PR 16-AUG-2001; 2001EP-0000374.
 PR 16-AUG-2001; 2001EP-0000377.
 PR 20-SEP-2001; 2001EP-0000478.

PR 20-SEP-2001; 2001EP-0000483.
 PR 22-OCT-2001; 2001EP-0000552.
 PR 22-OCT-2001; 2001EP-0000553.
 PR 22-OCT-2001; 2001EP-0000554.
 PR 22-OCT-2001; 2001EP-0000556.
 PR 22-OCT-2001; 2001EP-0000557.
 PR 22-OCT-2001; 2001EP-0000558.
 PR 15-NOV-2001; 2001EP-0004464.
 PR 21-DEC-2001; 2001EP-0005117.
 PA (STAM) DSM NV.
 XX
 XX Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
 PI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
 PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
 PI Stiebler J, Albarg R;
 XX
 DR WPI; 2002-723203/78.
 DR P-PSDB; ABZ78231, ABZ78288.
 XX
 PT Novel isolated protease polypeptide useful in laboratory, clinical,
 PT pharmaceutical, chemical, diagnostic, personal care and industrial
 PT applications
 XX
 XX Claim 13; Page 379-382; 394pp; English.
 XX
 CC The invention relates to a novel isolated protease polypeptide. A
 CC polypeptide or polynucleotide of the invention is useful for diagnosing a
 CC fungal infection such as aspergilliosis, or as a query sequence to perform
 CC a search against public databases. A polypeptide of the invention is
 CC useful in a selected number of industrial or pharmaceutical processes, in
 CC laboratory or clinical processes, in food industry (baking, brewing,
 CC cheese manufacturing, meat tenderising), in tanning industry and in the
 CC manufacture of biological detergents. A polypeptide may also be useful
 CC for improving protein solubility, extraction yields, viscosity or taste,
 CC texture, nutritional value, minimising of antigenicity or
 CC anti-nutritional factors, colour or functionality as well as processing
 CC aspects like filterability of the proteinaceous raw material. The
 CC sequences shown in ABR38313-ABR38869 represent the A. niger proteases of
 CC the invention.
 XX
 SQ Sequence 551 AA;
 Alignment Scores:
 Pred. No.: 1.39e-205 Length: 551
 Score: 2376.00 Matches: 428
 Percent Similarity: 87.96% Conservative: 54
 Best Local Similarity: 78.10% Mismatches: 66
 Query Match: 78.49% Indels: 0
 DB: 23 Gaps: 0
 US-09-712-338-1 (1-1669) x ABR38864 (1-551)
 QY 1 ATGCGTGGCTACGAATTTCTCTCAGTGTACCCCTGGTTCAGCCAGTGGGCCCTTCCA 60
 |||||
 Db 1 MetArgGlySerArgLeuValLeuLeuLeuProLeuAlaLeuSerCysAlaMetPro 20
 |||||
 QY 61 GGAAGTACACCGGCGCTCGGTAGAGACAGCTACCCCAAGAACCCACCGGGGTCAAG 120
 |||||
 Db 21 GluAsnGluTrpSerSerThrIleArgArgGlnLeuProLysAlaserThrGlyValLys 40
 |||||
 QY 121 ACTCTTACACCGCAACAATGTACCATCCGGTACAAAGAACCCGGGCGCAGGCGTC 180
 |||||
 Db 41 SerIleLysThrProAsnAsnValThrIleArgTyrLysGluProGlyThrGluGlyIle 60
 |||||
 QY 181 TGCAGAGCTACCCGGGTGTCAAATCTCTACTCTGGATATGCGACACCTCTCCGAGTCC 240
 |||||
 Db 61 CysGluThrThrProGlyValLysSerTyrSerGlyTyrValAspLeuSerProGluSer 80
 |||||
 QY 241 CATACCTTCTTCTGTTCTTTCGAAGCCAGACATACCCAGAAACTGCACCTATCATGTTG 300
 |||||
 Db 81 HisThrPhePheTrpPheGluSerArgArgAspProGluAsnAspProValThrLeu 100
 |||||

301 TGGTTGAATGGTGGCCCTGGAAGCGATTCTTTTGATCGCTCTCTTCGAAGAGTGGGCGCT 360
Db
101 TrpLeuasnGlyGlyProGlySerAspSerLeuileGlyLeuPheGluGluLeuGlyPro 120
Qy
361 TGCATGTCATTCGACTTTGATGATCAATCAACCTCTACTCGTGTGAAGAGGTCTCC 420
Db
121 CysHisIleThrProGluTyrGluSerIleIleAsnGlnTyrSerTrpAsnGluValThr 140
Qy
421 AATTACTATTCTGCTGCCAGCCATTGGAGTTCGGCTTTTCATATAGTATGATACGGTTGAT 480
Db
141 AsnLeuLeuPheLeuSerGlnProLeuGlyValGlyPheSerTyrSerGluThrGluAla 160
Qy
481 GGGTCCATTAACTTGAATCGGCTGCTCGAAATTCGAGCTTTTCAGGAGTTCAGGCG 540
Db
161 GlySerLeuAsnProPheThrGlyAlaValGluAsnAlaSerPheAlaGlyValGlnGly 180
Qy
541 CGGTACCAACCAATTGATGCTGCTGATCGATCACTACCAATCTTGGCCGAGAGCCGCT 600
Db
181 ArgTyrProValIleAspAlaThrIleIleAspThrThrAspIleAlaAlaArgAlaThr 200
Qy
601 TGGGAGATCTCGCAAGGATTCCTTAGTGACTACCTAGCTTGGACTCTAGGCTGCGAGTCT 660
Db
201 TrpGluValLeuGlnGlyPheLeuSerGlyLeuSerGlnLeuAspSerGluValIysSer 220
Qy
661 AAGGACTTCAGTCTATGACGAGAGCTATGAGGCGACTATGAGGCGACTATGCTCTCAT 720
Db
221 LysGluPheAsnLeuTrpThrGluSerTyrGlyGlyHisTyrGlyProAlaPheAsn 240
Qy
721 CATTTTACGACAGATGAGAGATTCGACGCTAGTGTAAATGGTCTTCAGCTTAAT 780
Db
241 HisPheTyrGluGlnAsnSerLysIleAlaSerGlyGluValAsnGlyValGlnLeuAsn 260
Qy
781 TTCAACTCTCTGGGAATTAATTAACGCTATCATCGAGCGGCTATCGAGCGCTTACTAC 840
Db
261 PheAsnSerLeuGlyIleIleAsnGlyIleIleAspAlaAlaIleGlnAlaAspTyrTyr 280
Qy
841 CTTGAATTCGCTGTGAACAATACCTACGCTATCAAGCTGTCAACGAGACCGTCTCAAC 900
Db
281 AlaAspPheAlaValAsnAsnThrTyrGlyIleLysAlaValAsnAspThrValTyrAsn 300
Qy
901 TCATGAGATTTCCCAACCAATGCCAATGGTTCGAGGATTTGATTTCCACCTCGAAA 960
Db
301 TyrMetLysPheAlaAsnThrMetProAsnGlyCysGlnAspGlnValAlaSerCysLys 320
Qy
961 CAGACAAACCGACCGCTATGACTAGCTAGCTCTGCGCCGAGAGCCCAACATGTGC 1020
Db
321 LeuThrAsnArgThrSerLeuSerAspTyrAlaIleCysThrGluAlaAlaAsnMetCys 340
Qy
1021 AGGCAATGTTGAGGGCCCATCTACGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db
341 ArgAspAsnValGluGlyProTyrTyrGlnPheGlyArgGlyValTyrAspIleArg 360
Qy
1081 CATCCATATGATGACCGCTCGCCGAGTATTAACAATAATTCCTGGCAAGAGTCTCT 1140
Db
361 HisProTyrAsnAspProThrProProSerTyrPheValAspTyrLeuLysAspSer 380
Qy
1141 GTCATGACCTATCGCGCTCAACATCACTACACCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db
381 ValMetAspAlaIleGlyValAspIleAsnTyrThrGluSerSerGlyGluValTyrTyr 400
Qy
1201 GCTTTCGAGAAACAGCGACTTGTGTGCGCCCACTTTCATCGAAGACCTCGAGGATC 1260
Db
401 AlaPheGlnGlnThrGlyAspPheValTrpProAsnPheIleGluAspLeuGluIle 420
Qy
1261 CTTTCTCTCCCGGCTGCT 1320
Db
421 LeuGlnLeuProValArgValSerLeuIleIleIleIleIleIleIleIleIleIleIle 440
Qy
1321 TTGGCGGTGAGCGCGTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Db
441 PheGlyGlyGlnAlaIleSerLeuAlaValAsnTyrProHisAlaAlaGlnPheArgAla 460
Qy
1381 GCAGGGTACAGCCCTCGAAAGTCAACGGCTCGAGTATGGGGAACCTCGCGAGTATGGT 1440

461 AlaGlyTyrThrProMetThrValAspGlyValGluTyrGlyGluThrArgGluTyrGly 480
1441 AATTTCTCTTCACTCGCGTCTATGAGCAGCGCATGAAGTCCCATACTACAGCCCATC 1500
481 AsnPheSerPheThrArgValTyrGlnAlaGlyHisGluValProTyrTyrGlnProIle 500
1501 GCCTCCTGCAATGTTTAAACCGGACTATCTTCGTTGGGATATCGAGAGGCCAGAG 1560
501 AlaAlaLeuGlnLeuPheAsnArgThrLeuPheGlyTrpAspIleAlaAlaGlyThrThr 520
1561 AAGATCTCGCCGCTACAGACCAATGGAACGGCTACAGCTACGATACACAGTCTGCTCC 1620
521 GlnIleTrpProGluTyrSerThrAsnGlyThrSerGlnAlaThrHisThrGluSerPhe 540
1621 GTGCGCTGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTAC 1640
541 ValProLeuSerThrAlaSerSer 548

RESULT 3
ABR38865
ID ABR38865 standard; Protein; 623 AA.
XX ABR38865;
AC ABR38865;
XX 24-APR-2003 (first entry)
DT
DE A. niger serine carboxypeptidase polypeptide #4.
KW Protease; fungal infection; aspergillosis; food; tanning; detergent;
KW protein solubility; viscosity; taste; texture; nutritional value;
KW EC3.4.16.6.
XX Aspergillus niger.
OS
PN WO200258623-A2.
PD
PF 06-SEP-2002.
XX 22-FEB-2002; 2002WO-EP01984.
XX 23-FEB-2001; 2001EP-0200657.
PR 23-FEB-2001; 2001EP-0200658.
PR 23-FEB-2001; 2001EP-0200660.
PR 26-FEB-2001; 2001EP-0200706.
PR 26-FEB-2001; 2001EP-0200707.
PR 26-FEB-2001; 2001EP-0200708.
PR 26-FEB-2001; 2001EP-0200719.
PR 28-MAR-2001; 2001EP-0000075.
PR 28-MAR-2001; 2001EP-0000078.
PR 28-MAR-2001; 2001EP-0000080.
PR 28-MAR-2001; 2001EP-0000087.
PR 28-MAR-2001; 2001EP-0000088.
PR 21-MAY-2001; 2001EP-0000156.
PR 21-MAY-2001; 2001EP-0000159.
PR 21-MAY-2001; 2001EP-0000160.
PR 21-MAY-2001; 2001EP-0000162.
PR 21-MAY-2001; 2001EP-0000165.
PR 21-MAY-2001; 2001EP-0000166.
PR 21-MAY-2001; 2001EP-0000168.
PR 21-JUN-2001; 2001EP-0000168.
PR 21-JUN-2001; 2001EP-0000240.
PR 21-JUN-2001; 2001EP-0000242.
PR 21-JUN-2001; 2001EP-0000244.
PR 21-JUN-2001; 2001EP-0000246.
PR 12-JUL-2001; 2001EP-0000280.
PR 12-JUL-2001; 2001EP-0000285.
PR 30-JUL-2001; 2001EP-0000323.
PR 30-JUL-2001; 2001EP-0000327.
PR 02-AUG-2001; 2001EP-0000341.
PR 02-AUG-2001; 2001EP-0000342.
PR 02-AUG-2001; 2001EP-0000343.
PR 02-AUG-2001; 2001EP-0000344.
PR 09-AUG-2001; 2001EP-0000357.

```

PR 16-AUG-2001; 2001EP-0000374.
PR 16-AUG-2001; 2001EP-0000377.
PR 20-SEP-2001; 2001EP-0000478.
PR 20-SEP-2001; 2001EP-0000483.
PR 22-OCT-2001; 2001EP-0000552.
PR 22-OCT-2001; 2001EP-0000553.
PR 22-OCT-2001; 2001EP-0000554.
PR 22-OCT-2001; 2001EP-0000556.
PR 22-OCT-2001; 2001EP-0000557.
PR 22-OCT-2001; 2001EP-0000558.
PR 15-NOV-2001; 2001EP-0004464.
PR 21-DEC-2001; 2001EP-0005117.
PA (STAM ) DSM NV.
XX
XX Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
PI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
XX Stiehler J, Albarg R;
DR
DR WPI; 2002-723203/78.
DR P-PSTD; ABZ78232, ABZ78289.
XX
XX Novel isolated protease polypeptide useful in laboratory, clinical,
PT pharmaceutical, chemical, diagnostic, personal care and industrial
PT applications -
XX
PS Claim 13; Page 382-385; 394pp; English.
XX
CC The invention relates to a novel isolated protease polypeptide. A
CC polypeptide or polynucleotide of the invention is useful for diagnosing a
CC fungal infection such as aspergillosis, or as a query sequence to perform
CC a search against public databases. A polypeptide of the invention is
CC useful in a selected number of industrial or pharmaceutical processes, in
CC laboratory or clinical processes, in food industry (baking, brewing,
CC cheese manufacturing, meat tenderising), in tanning industry and in the
CC manufacture of biological detergents. A polypeptide may also be useful
CC for improving protein solubility, extraction yields, viscosity or taste,
CC texture, nutritional value, minimising of antigenicity or
CC anti-nutritional factors, colour or functionality as well as processing
CC aspects like filterability of the proteinaceous raw material. The
CC sequences shown in ABR38313-ABR38869 represent the A. niger proteases of
CC the invention.
XX
SQ Sequence 623 AA;
Alignment Scores:
Pred. No.: 1.66e-79 Length: 623
Score: 980.50 Matches: 216
Percent Similarity: 56.23% Conservative: 91
Best Local Similarity: 39.56% Mismatches: 178
Query Match: 32.39% Indels: 61
DB: 23 Gaps: 19
US-09-712-338-1 (1-1668) x ABR38865 (1-623)
QY 130 ACCGCAACAAATGTACCATTCCGGTACAAGGA-----CCGGGGCAGAG----- 174
||| :|||||: :||| :|||||
Db 32 ThrProGluAspLeuThrValIleHisSerGluIlePheProGlyAlaArgIleSerTyr 51
QY 175 -----GCCGTCGGAGACTACCCCGGGTGTCAAATCCTACTCTGGATATGC 222
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db 52 LysGlnProLeuGlyIleCysThrThrProSerThrProSerTyrSerGlyTyrIle 71
QY 223 GACACCTCTCCGAGTCC-----:|||: 240
|||: |||:
Db 72 HisLeuProProHisThrLeuThrAsnLeuSerIleProGlyIleSerIleSerGlnPro 91
QY 241 -----CATACCTTCTTCTGTTCTTCCGAAGCCAGA-----CATATCCCAGAACT 285
:|||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db 92 TyrProIleAsnThrPhePheTrpTyPheProSerArgHisHisAsnAspThr 111
QY 286 GCACCTATCACATGTGGTGAATGGTGGCCCTGGAAGCGATTCTTTGATCGGTCTCTC 345

```

Db 459 ArgGlyGlyGluAspValSerLeuValGluTyrGluAspAlaGluLysPheArgAla 478
Qy 1381 GCAGGTACACGCCCTCAAGTCAACGGCGTCGAGTAT-----GGGAAACTCGCGAG 1434
Db 479 AlaGlyTyrAlaGluValGlnThrLysSer---SeryrValGlyGlyLeuValArgGln 497
Qy 1435 TATGTAATTTCTCTTCACTCGGCTCTATGAGCGAGCCCATGAGTCCCATACACAG 1494
Db 498 TyrGlyAsnPheSerPheThrArgValPheGlnAlaGlyHisGluValProPheTyrGln 517
Qy 1495 CCCATCGGCTCCCTGCAATGTTAAACCGGACTATCTTCGGTTGGATATCGCAGAGGC 1554
Db 518 ProGluThrAlaTyrGluLeuPheAsnArgAlaGlnPheAsnTrpAspIleAlaThrGly 537
Qy 1555 CAGAAGAGATCGG-----CCGAGCTACAGACGAATGGAAACGCTACAGCTACGCAT 1608
Db 538 GlyIleSerLeuGluGlnAsnGlnSerTyrGlyThrGluGlyProSerSerThrTrpHis 557
Qy 1609 ACACAGTCTCGTGGCGG 1626
Db 558 IleLysAsnGluValPro 563
RESULT 4
ABR38859
ID ABR38859 standard; Protein; 554 AA.
XX ABR38859;
DT
DT 24-APR-2003 (first entry)
XX
DE A. niger serine carboxypeptidase polypeptide #2.
XX
XX Protease; fungal infection; aspergillosis; food; tanning; detergent;
KW protein solubility; viscosity; taste; texture; nutritional value;
KW EC3.4.16.6.
OS Aspergillus niger.
XX
XX W0200268623-A2.
PN
XX
XX
XX 06-SEP-2002.
PD
XX
XX 22-FEB-2002; 2002W0-EP01984.
PR
XX 23-FEB-2001; 2001EP-0200657.
PR 23-FEB-2001; 2001EP-0200658.
PR 23-FEB-2001; 2001EP-0200660.
PR 26-FEB-2001; 2001EP-0200706.
PR 26-FEB-2001; 2001EP-0200707.
PR 26-FEB-2001; 2001EP-0200708.
PR 26-FEB-2001; 2001EP-0200719.
PR 28-MAR-2001; 2001EP-0000075.
PR 28-MAR-2001; 2001EP-0000078.
PR 28-MAR-2001; 2001EP-0000080.
PR 28-MAR-2001; 2001EP-0000087.
PR 28-MAR-2001; 2001EP-0000088.
PR 21-MAY-2001; 2001EP-0000156.
PR 21-MAY-2001; 2001EP-0000159.
PR 21-MAY-2001; 2001EP-0000160.
PR 21-MAY-2001; 2001EP-0000162.
PR 21-MAY-2001; 2001EP-0000165.
PR 21-MAY-2001; 2001EP-0000166.
PR 21-MAY-2001; 2001EP-0000168.
PR 21-JUN-2001; 2001EP-0000240.
PR 21-JUN-2001; 2001EP-0000242.
PR 21-JUN-2001; 2001EP-0000244.
PR 21-JUN-2001; 2001EP-0000246.
PR 12-JUL-2001; 2001EP-0000285.
PR 12-JUL-2001; 2001EP-0000287.
PR 30-JUL-2001; 2001EP-0000323.
PR 30-JUL-2001; 2001EP-0000327.
PR 02-AUG-2001; 2001EP-0000341.
PR 02-AUG-2001; 2001EP-0000342.

PR 02-AUG-2001; 2001EP-0000343.
PR 02-AUG-2001; 2001EP-0000344.
PR 09-AUG-2001; 2001EP-0000357.
PR 16-AUG-2001; 2001EP-0000374.
PR 16-AUG-2001; 2001EP-0000377.
PR 20-SEP-2001; 2001EP-0000478.
PR 20-SEP-2001; 2001EP-0000483.
PR 22-OCT-2001; 2001EP-0000552.
PR 22-OCT-2001; 2001EP-0000553.
PR 22-OCT-2001; 2001EP-0000554.
PR 22-OCT-2001; 2001EP-0000556.
PR 22-OCT-2001; 2001EP-0000557.
PR 22-OCT-2001; 2001EP-0000558.
PR 15-NOV-2001; 2001EP-0004464.
PR 21-DEC-2001; 2001EP-0005117.
XX
XX (STAM) DSM NV.
PA Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
PI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
PI Stiebeler J, Albarg R;
XX
XX WPI: 2002-723203/78.
DR P-PSDB; ABZ78226, ABZ78283.
DR
XX
XX Novel isolated protease polypeptide useful in laboratory, clinical,
PT pharmaceutical, chemical, diagnostic, personal care and industrial
PT applications -
XX
XX Claim 13; Page 366-369; 394pp; English.
PS
XX The invention relates to a novel isolated protease polypeptide. A
CC polypeptide or polynucleotide of the invention is useful for diagnosing a
CC fungal infection such as aspergillosis, or as a query sequence to perform
CC a search against public databases. A polypeptide of the invention is
CC useful in a selected number of industrial or pharmaceutical processes, in
CC laboratory or clinical processes, in food industry (baking, brewing,
CC cheese manufacturing, meat tenderising), in tanning industry and in the
CC manufacture of biological detergents. A polypeptide may also be useful
CC for improving protein solubility, extraction yields, viscosity or taste,
CC texture, nutritional value, minimising of antigenicity or
CC anti-nutritional factors, colour or functionality as well as processing
CC aspects like filterability of the proteinaceous raw material. The
CC sequences shown in ABR38313-ABR38889 represent the A. niger proteases of
CC the invention.
XX
XX SQ Sequence 554 AA;
XX
XX Alignment Scores:
Pred. No.: 1.67e-76 Length: 554
Score: 947.00 Matches: 215
Percent Similarity: 51.15% Conservatives: 73
Best Local Similarity: 38.19% Mismatches: 176
Query Match: 31.29% Indels: 99
DB: 23 Gaps: 14
US-09-712-338-1 (1-1668) x ABR38859 (1-554)
Qy 91 CAGCTACCCCAAGAACCCACCGGGGTCAAGACTTTTACAACCGCA-----ACAATGTC 144
Db 17 GlnPheProGluProGluGlyIleThrValLeuLysSerLysLeuHisGluAsnVal 36
Qy 145 ACCATCGGTACAGGACCGGGGCGAGAGCTCTCCGAGTCCCATACCTCTTCTGTTCTTCGAA 204
Db 37 ThrIleSerPheLysGluPro-----GlyIleCysGluThrThrProGlyValArg 53
Qy 205 TCCTACTCTGGATATGTCACACACCTCTCCCGAGTCCCATACCTCTTCTGTTCTTCGAA 264
Db 54 SerTyrSerGlyTyrValHisLeuProAlaSerThrSerPhePheThrPhePheGlu 73
Qy 265 GCCAGACATAACCCAGAAACTGCACCTATCATCTGTTGATGTTGGCCCTGGAAGC 324

QY	1321	TTCGGCGGTCAAGCCGTTTCCCTCGCTGCAGAACTACTCCCAAGCCGCCAGTCCGAAGC	1380
Db	397	ValGlyGlyGluLysAlaSerLeuAlaValProTy:SerArgIleThrGluPheAlaAsp	416
QY	1381	GCAGGTACAGCCCTCAAAGTCAACGGCGTCGAGTATGGGAAAACCTCGCAGTAGTGCT	1440
Db	417	ThrGlyTyrSerProLeuLeuThrProAspGlyIleSerGlyMetThrArgGlnLeuGly	436
QY	1441	AATTTCCTCCCTCACTCGCGTCTATGAGCGAGGCCATCAAGTCCCATCTACCAGCCCACATC	1500
Db	437	AsnTyrSerPheThrArgValPheGlnAlaGlyHisGluValProSerTyrGlnProVal	456
QY	1501	GCTTCCTCGAATGTTTAACCGAGCTATCTTCGGTTGGATATCCAGAGGCCACAAG	1560
Db	457	AlaAlaTyrGluIlePheMetAlaGlnThrPheAsnLysAspileProThrGlyLeuLeu	476
QY	1561	AAGATCTGCCCGCCAGCTACAGACGAATGGACGGCTACAGCTACGATACACAGTCGTCC	1620
Db	477	AlaValAspAspGluPheGlnSerValGlyProLysAspThrTrpHisIleLysAsnIle	496
QY	1621	GTGCGCGTCCCTACGGCTACCGACATGTCAGAGTGTG-----	1657
Db	497	ProProile-MetProLysProGlnCysTyrValLeuSerProGlyThrCysThrProGl	516
QY	1658	-GTATGG	1663
Db	516	uValTrp	518
RESULT 5			
ID	ABR38817		
XX	ABR38817 standard; Protein; 526 AA.		
AC	ABR38817;		
DT	24-APR-2003 (first entry)		
DE	A. niger serine carboxypeptidase polypeptide #1.		
KW	Protease; fungal infection; aspergillosis; food; tanning; detergent;		
KW	protein solubility; viscosity; taste; texture; nutritional value;		
KW	EC3.4.16.6.		
OS	Aspergillus niger.		
PN	WO20026823-A2.		
PD	06-SEP-2002.		
PF	22-FEB-2002; 2002WO-EPO1984.		
PR	23-FEB-2001; 2001EP-0200657.		
PR	23-FEB-2001; 2001EP-0200658.		
PR	23-FEB-2001; 2001EP-0200660.		
PR	26-FEB-2001; 2001EP-0200706.		
PR	26-FEB-2001; 2001EP-0200707.		
PR	26-FEB-2001; 2001EP-0200708.		
PR	26-FEB-2001; 2001EP-0200719.		
PR	28-MAR-2001; 2001EP-0000075.		
PR	28-MAR-2001; 2001EP-0000078.		
PR	28-MAR-2001; 2001EP-0000080.		
PR	28-MAR-2001; 2001EP-0000087.		
PR	28-MAR-2001; 2001EP-0000088.		
PR	21-MAY-2001; 2001EP-0000156.		
PR	21-MAY-2001; 2001EP-0000159.		
PR	21-MAY-2001; 2001EP-0000160.		
PR	21-MAY-2001; 2001EP-0000162.		
PR	21-MAY-2001; 2001EP-0000165.		
PR	21-MAY-2001; 2001EP-0000166.		
PR	21-MAY-2001; 2001EP-0000168.		
PR	21-JUN-2001; 2001EP-0000240.		
PR	21-JUN-2001; 2001EP-0000242.		
PR	21-JUN-2001; 2001EP-0000244.		

PR 21-JUN-2001; 2001EP-0000246.
 PR 12-JUL-2001; 2001EP-0000280.
 PR 12-JUL-2001; 2001EP-0000285.
 PR 30-JUL-2001; 2001EP-0000323.
 PR 30-JUL-2001; 2001EP-0000327.
 PR 02-AUG-2001; 2001EP-0000341.
 PR 02-AUG-2001; 2001EP-0000342.
 PR 02-AUG-2001; 2001EP-0000343.
 PR 02-AUG-2001; 2001EP-0000344.
 PR 02-AUG-2001; 2001EP-0000357.
 PR 16-AUG-2001; 2001EP-0000374.
 PR 16-AUG-2001; 2001EP-0000377.
 PR 20-SEP-2001; 2001EP-0000478.
 PR 20-SEP-2001; 2001EP-0000483.
 PR 22-OCT-2001; 2001EP-0000552.
 PR 22-OCT-2001; 2001EP-0000553.
 PR 22-OCT-2001; 2001EP-0000554.
 PR 22-OCT-2001; 2001EP-0000556.
 PR 22-OCT-2001; 2001EP-0000557.
 PR 22-OCT-2001; 2001EP-0000558.
 PR 15-NOV-2001; 2001EP-0004464.
 PR 21-DEC-2001; 2001EP-0005117.

(STAM) DSM NV.

XX Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
 PI Klugbauer S, Wagner C, Fritz A, Von Gusted W, Heinrich O;
 PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
 PI Stiebler J, Albarg R;

XX WPI; 2002-723203/78.
 DR P-PSDB; ABZ78184, ABZ78241.

XX Novel isolated protease polypeptide useful in laboratory, clinical,
 PT pharmaceutical, chemical, diagnostic, personal care and industrial
 PT applications

XX Claim 13; Page 271-274; 394pp; English.

XX The invention relates to a novel isolated protease polypeptide. A
 CC polypeptide or polynucleotide of the invention is useful for diagnosing a
 CC fungal infection such as aspergillosis, or as a query sequence to perform
 CC a search against public databases. A polypeptide of the invention is
 CC useful in a selected number of industrial or pharmaceutical processes, in
 CC laboratory or clinical processes, in food industry (baking, brewing,
 CC cheese manufacture, meat tenderising), in tanning industry and in the
 CC manufacture of biological detergents. A polypeptide may also be useful
 CC for improving protein solubility, extraction yields, viscosity or taste,
 CC texture, nutritional value, minimising of antigenicity or
 CC anti-nutritional factors, colour or functionality as well as processing
 CC aspects like filterability of the proteinaceous raw material. The
 CC sequences shown in ABR38313-ABR38869 represent the A. niger proteases of
 CC the invention.

XX Sequence 526 AA;

Alignment Scores:
 Pred. No.: 1.58e-74 Length: 526
 Score: 925.00 Matches: 221
 Percent Similarity: 51.46% Conservative: 78
 Best Local Similarity: 38.04% Mismatches: 192
 Query Match: 30.56% Indels: 90
 DB: 23 Gaps: 15

US-09-712-338-1 (1-1668) x ABR38817 (1-526)

QY 10 TACCAATTTCTCAGTCCTACCTTGGTTCAGCCAGTGGCGCTTCCAGGAGTACA 69
 Db 2 TyrTyrSerLeuTrpValAlaLeuValAla-----AlaLeuPro----- 15
 QY 70 CCGGCGTCCGTCGGTAGAAGACAGCTACCCCAAGAACCCCGGGGTCAAGACTCTTACA 129
 Db 16 -----ValSerArgAlaGlnPheValAlaProProThrAspLeuIleProThrlys 32

QY 130 ACCGCAACAATGTCACCATCGGTACAGAACCCGGGGCAGAGGGGCTTCGCGAGCT 189
 Db 33 GlyTyrLeuAspIleProValArgTyrLysGlnVal---ProThrGlyIleCysGlnThr 51
 QY 190 ACCCGGGTGTCAAATCCTACTCTCGATATGCGACACCTCTCCGAGTCCCATACCTTC 249
 Db 52 AspProSerValLysSerPheSerGlyTyrValAspValAlaGluHisGluHisIlePhe 71
 QY 250 TTCTGTTCTTCGAAGCCAGACAT---AACCCAGAAACTGCACCTATCATCTGTTGGTTG 306
 Db 72 PheTrpPheGluAlaArgAsnGlnAspProThrGluAlaProLeuThrValTrpIle 91
 QY 307 AATGGT-----GGCCTCGAAGCGATTCTTTCATCGGTCTCTTCAAGAG 351
 Db 92 AsnGlyGlyMetSerAspProGlyProGlySerSerMetIleGlyLeuGlnGlu 111
 QY 352 TTGGGCGCTTGCATGCAATTCGACTTTTGTGACTACATCAACCTCCTCGTGAAC 411
 Db 112 HisGlyProCysGlyIleAspAlaAsnGlySerValTyrAsnAsnProTyrSerTrpAsn 131
 QY 412 GAGGTCTCAAATTTACTATCTCTGTCGCCAGCATCGGAGTCGGCTTTTCATATAGTAT 471
 Db 132 AsnAlaSerAsnMetLeuTyrIleAspGlnProValGlnThrGlyPheSerTyrSer--- 150
 QY 472 ACGTGTGATGGTCCATTAACTCTGTAACCTGTAACTGGGTCTCGAAATTCGAGCTTGCAGGA 531
 Db 151 -----lleProValProGlyTyrValAspSerSerThrAspAsnGly 164
 QY 532 GTTCAGGGCGGTACCCCAACCATTTGATGCCACTCTGTATCATCACTACCAATCTTGGCGCA 591
 Db 165 PheMetGlyAlaPhePro----- 170
 QY 592 GAGGCGCTTGGGAGATCTCTGCAAGATTCTCTTAGTGGACTACTAGCTTGGACTTAGG 651
 Db 171 -----Gln 171
 QY 652 GTCCAGCTTAAGACTTCAGTCTATGGACGGAGAGCTATGGAGGGCACTATGTCTCTGCA 711
 Db 172 TyrSerArgGlnThrPheHisPheThrThrGluSerTyrGlyGlyHisTyrGlyProVal 191
 QY 712 TTCTTCAATCATTTTACGAGCAGAGATGAGAGAATTCGCCAAGGTAGTGTAAATGTGTT 771
 Db 192 PheAsnGluTyrIleGluGlnAsnAlaHisLeuGlnProGlyAla-----Lys 208
 QY 772 CAGCTTAATTTCAACTCTCTGGGAATATTAAACGGCATCATCAGCAGGGGATCCAGGCC 831
 Db 209 LysIleGlnLeuGlySerValMetIleGlyAsnGlyTrpTyrAspProIleGlnTyr 228
 QY 832 CCTTACTACCTGAATTCGCTGTG-----AACAAATACCTACGGTATCAAGGCTGTCAAC 885
 Db 229 GlnAlaTyrTyrAsnPheThrValTyrProGlyAsnThrTyrAspTyrLeuProPheAsn 248
 QY 886 GAGACCGCTTACAACTACATGAAGTTTGGCCAAACAAATGCCAATGGTTGCCAGGATTG 945
 Db 249 LysSerIleSerSerLeuMetTyrAsnAsnLeuTyrGlyProGlyAsnCysLeuAspGln 268
 QY 946 ATTTCCACCTGCAAAACAGAACCCGACCGCATAGCTAGCTAGCCCTCTGCGCCGAA 1005
 Db 269 LeuTyrAspCys-----AlaAlaArgGlyIleAspGluIleCysSerThr 283
 QY 1006 GCCACCAACATGTGCAGGGACAATGTTGAGGGGCCATACCTAGCGCTTGTGGTCTGGTGT 1065
 Db 284 AlaAspAspPheCysAlaAsnGluValGluAsnValTyrAspIleTyrSerGlyArgAsp 303
 QY 1066 GTGTATGATATTCGGCATCCATATGATGACCCGACTCCGCCAAGTTATTACACAAATTT 1125
 Db 304 GluTyrAspPheArgGluLeuThrProAspProPheProTyrGluPheTyrValAspTyr 323
 QY 1126 CTGCAAGAGGACTCTGTCTAGGAGCTATCGGCTCAACATCACTCACTACCCAGTCCAAT 1185
 Db 324 LeuAsnLysAlaSerValGlnAlaAlaIleGlyAlaTyrIleAsnTyrThrGluSerAsn 343

QY 523 TTTCAGGAGTTCAGGCGCGTACCCACCAATTGATGCCACTCTGATCGATACCTACCAAT 582
 Db 182 -----GlulysValSerThrLys 188
 QY 583 CTTGCCGAGAGCGCTTGGGAGATCTTCAAGATTCCTTAGTGGACTACCTAGCTTG 642
 Db 189 LeuAlaGlyLysAspAlaTyrIlePheLeuGluLeuPheGluAlaPheProHisLeu 208
 QY 643 GACTCTAGGCTCAGTCTAAGACTTCACTGATGACGAGAGCTATGAGGGGCACTAT 702
 Db 209 -----ArgSerAsnAspPheHisIleAlaGlySerTyrAlaGlyHisTyr 224
 QY 703 GCTCTGCTCATCTTCAATCATTTTACGAGCAGAT---GAGAGAAATGCCAACGGTAGT 759
 Db 225 IleProGlnIleAlaHisGluIleValValLysAsnProGluArg----- 239
 QY 760 GTTATGGTGTTCAGCTTAATTCATCTCTGGGAATTTAAGCGCATCATCGACGAG 819
 Db 240 -----ThrPheAsnLeuThrSerValMetIleGlyAsnGlyIleThrAspPro 255
 QY 820 GCGATCCAGGCGCTTACTACCTGAAATTCGCTGTG---AACAAATACCTACGGTATCAAG 876
 Db 256 LeuIleGlnAlaAspTyrTyrGluProMetAlaCysGlyLysGlyTyrHisProVal 275
 QY 877 GCTGTCAACGAGACCGCTTACAACTACATGAGTTTGGCAACCAATGCCAATGTTGC 936
 Db 276 LeuSerSerGluGluCysGluLysMetSerLysAlaAlaGlyArg----- 290
 QY 937 CAGGATTTGATTTCCACCTGCAACAGACACACCGCAGCCGATTCAGTACGACGCTC 996
 Db 291 -----CysArgArgLeuAsnLysLeuCysTyrAlaSerLysSerSer 304
 QY 997 -----TGGCGCCAGCCCAACATGTGCGAGGCAATTTGAGGGGCGCATACTACGCC 1050
 Db 305 LeuProCysIleValAlaThrAlaTyrCysAspSerAlaLeuLeuGluProTyrIleAsn 324
 QY 1051 TTTCGTGGTGGTGTATGATATTCGGCATCATCATATGATGACCGG----- 1098
 Db 325 ---ThrGlyLeuAsnValTyrAspIleArgGlyProCysGluAspAsnSerThrAspGly 343
 QY 1099 -----ACTCGGCCAGTTTATACAAATTTCTGCAAGGACTCTGTGTCATGGAC 1149
 Db 344 MetCysTyrThrGlyLeuArgTyrValAspGlnTyrMetAsnPheProGluValGlnGlu 363
 QY 1150 GCTATCGGCTCAACATC---AACTACACCCAGCTCCCAATATACGCTTCTACTACGTTTC 1206
 Db 364 ThrLeuGlySerAspValHisAsnTyrSerGlyCysAspAsnAspValPheThrGlyPhe 383
 QY 1207 CAGCAACAGGCGACTTGTCTGGCCCACTTCATCGAACCTCGAGGAGATCCTTGCT 1266
 Db 384 LeuPheThrGlyAspGlySerLysPro---PheGlnGlnTyrIleAlaGluLeuAsn 402
 QY 1267 CTCCCGTGGTGTCTCCCTCATCTATGCGGCGCGGATTCATCTCACTGCTTTCGCG 1326
 Db 403 HisAsnIleProValLeuIleTyrAlaGlyAspLysAspTyrIleCysAsnTrpLeuGly 422
 QY 1327 GGTGAGCGGCTTTCCTCGTGGCACTACTCCCAAGCCCGCCAGTTCGGAAGCGCAGGG 1386
 Db 423 AsnHisAlaTrpSerAsnGluLeuGluTrpIleAsnLysArgArgTyrGlnArgMet 442
 QY 1387 TACAGGCC-----CTGAAGTCAAGCGGTCTCAGTATGGGAACTCGCGAGTATGGT 1440
 Db 443 LeuArgProTrpValSerLysGluThrGlyGluGluLeuGlyGlnValLysAsnTyrGly 462
 QY 1441 AATTTCCTTCACTCGCTGTATGAGCGAGCGCATGAAGTCCCATCTACACGCCCATC 1500
 Db 463 ProPheThrPheLeuArgIleTyrAspAlaGlyHisMetValProTyrAspGlnProGlu 482
 QY 1501 GCCTCCCTGCAATGTTTAAACCGGACTATCTTCGGT 1536
 Db 483 AlaSerLeuGluMetValAsnSerTrpIleSerGly 494

RESULT 7

ABR38819
 ID ABR38819 standard; Protein; 536 AA.
 XX
 AC ABR38819;
 DT 24-APR-2003 (first entry)
 XX
 DE A. niger carboxypeptidase Y polypeptide #1.
 XX
 KW protease; fungal infection; aspergillosis; food; tanning; detergent;
 KW protein solubility; viscosity; taste; texture; nutritional value;
 KW EC3.4.16.5.
 XX
 OS Aspergillus niger.
 XX
 PN WO200268623-A2.
 XX
 PD 06-SEP-2002.
 XX
 PF 22-FEB-2002; 2002WO-EP01984.
 XX
 PR 23-FEB-2001; 2001EP-0200657.
 PR 23-FEB-2001; 2001EP-0200658.
 PR 23-FEB-2001; 2001EP-0200660.
 PR 26-FEB-2001; 2001EP-0200706.
 PR 26-FEB-2001; 2001EP-0200707.
 PR 26-FEB-2001; 2001EP-0200708.
 PR 26-FEB-2001; 2001EP-0200719.
 PR 28-MAR-2001; 2001EP-0300075.
 PR 28-MAR-2001; 2001EP-0300078.
 PR 28-MAR-2001; 2001EP-0300080.
 PR 28-MAR-2001; 2001EP-0300087.
 PR 28-MAR-2001; 2001EP-0300088.
 PR 21-MAY-2001; 2001EP-0300156.
 PR 21-MAY-2001; 2001EP-0300159.
 PR 21-MAY-2001; 2001EP-0300160.
 PR 21-MAY-2001; 2001EP-0300162.
 PR 21-MAY-2001; 2001EP-0300165.
 PR 21-MAY-2001; 2001EP-0300166.
 PR 21-MAY-2001; 2001EP-0300168.
 PR 21-JUN-2001; 2001EP-0300240.
 PR 21-JUN-2001; 2001EP-0300242.
 PR 21-JUN-2001; 2001EP-0300244.
 PR 21-JUN-2001; 2001EP-0300246.
 PR 12-JUL-2001; 2001EP-0300280.
 PR 12-JUL-2001; 2001EP-0300285.
 PR 30-JUL-2001; 2001EP-0300323.
 PR 30-JUL-2001; 2001EP-0300327.
 PR 02-AUG-2001; 2001EP-0300341.
 PR 02-AUG-2001; 2001EP-0300342.
 PR 02-AUG-2001; 2001EP-0300343.
 PR 02-AUG-2001; 2001EP-0300344.
 PR 09-AUG-2001; 2001EP-0300357.
 PR 16-AUG-2001; 2001EP-0300374.
 PR 16-AUG-2001; 2001EP-0300377.
 PR 20-SEP-2001; 2001EP-0300478.
 PR 20-SEP-2001; 2001EP-0300483.
 PR 22-OCT-2001; 2001EP-0300552.
 PR 22-OCT-2001; 2001EP-0300553.
 PR 22-OCT-2001; 2001EP-0300554.
 PR 22-OCT-2001; 2001EP-0300556.
 PR 22-OCT-2001; 2001EP-0300557.
 PR 22-OCT-2001; 2001EP-0300558.
 PR 15-NOV-2001; 2001EP-0300464.
 PR 21-DEC-2001; 2001EP-03005117.
 XX
 PA (STAM) DSM NV.
 XX
 EI Edens L, Van Dijk AA, Krubasik P, Albertmann K, Stock A, Kimpel E;
 EI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
 EI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
 EI Stiebler J, Albarg R;
 XX

Db 254 HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu 273
QY 730 GAGCAAGAATGAGAGAATATGCCAAACGGTAGCTGTAATGGTGTTCAGCTTAATTTCAACTCT 789
Db 274 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 284
QY 790 CTGGGAATATTACCGGATCATCGACGAGCGATCCAGGCCCTTACTACCTCGAATTC 849
Db 285 ValLeuIleGlyAsnGlyLeuThrAspGlyTyrThrGlnTyrGluTyrTyrArgProMet 304
QY 850 GCTGTGAACAATACCTACGGTATCAAGGCTCTCAACGAGACCGCTCTACAACATCATGAAG 909
Db 305 AlaCysGlyAspGlyGlyTyrProAlaValAsnAspGluSerSerCysGlnSerMetAsp 324
QY 910 TTTGCCAACCAATATGCCAAATGGTTGCCAGGATTTGATTTCCACTGCMAACACACACAAC 969
Db 325 -----AsnAlaLeuProArg-----CysGlnSerMetIleGluSerCysTyrSerSerGlu 341
QY 970 CGCACCGCATAGCTAGCTACGCCCTCTGGCGGAGCCACCAACATGTGCGAGGCAAT 1029
Db 342 SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAsnAla 356
QY 1030 GTTCAGGGGCCATACTACGCTTTGCTGGTTCGTGTGTATGATATTCGGCATCCATAT 1089
Db 357 LeuLeuAlaProTyrGlnArg---ThrGlyGlnAsnValTyrAspValArgGlyLysCys 375
QY 1090 GATGACCCG-----ACTCCGCCAAGTTATTACAAACAATTTCTGGCAAG 1134
Db 376 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 395
QY 1135 GACTCTCTCATGAGCGTATCGGGGTCAACATCAAC---TACACCCAGTCCCAATAATGAC 1191
Db 396 ProGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnPheAsp 415
QY 1192 GTCTACTACTCTTCCAGCAACAGCGGCTTTGCTCGGCC-----AACTTCATCGAA 1245
Db 416 IleAsnArgAsnPheLeuPheHisGlyAspTrpMetLysProTyrHisArgLeuValPro 435
QY 1246 GACCTCGAGGAGATCTTGTCTCTCCCGTGGTGTCTCCCTCATCTAT---GGCGAGCC 1302
Db 436 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 450
QY 1303 GATTACATCTGCAACTGTTCGGCGGTTCAGGCGGTTCCTCGCTCGAAGTACTCCCAA 1362
Db 451 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTyrProGly 470
QY 1363 GCGCGCCAGTTCGGAAGCGGAGGTACACGCCCTCGAAAGTC-----AAC 1407
Db 471 GlnAlaGluTyrAlaSerAlaGluLeuGluAspLeuValIleValAspAsnGluHisThr 490
QY 1408 GCGCTCGAGTATGGGAAACTCGCAGTAGTAGTAATTTCTCTCTCACTCGCGTCTATGAG 1467
Db 491 GlyLysLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 510
QY 1468 GCAGGCCATCAAGTCCCATCTACAGGCCCATCGGCTCCCTGCAATGTTTAAACCGGACT 1527
Db 511 GlyGlyHisMetValProMetAspGlnProGluSerSerLeuGluPhePheAsnArgTrp 530
QY 1528 ATCTCCGCT---TGG 1539
Db 531 LeuGlyGlyGluTrp 535

RESULT 8
AAR96737
ID AAR96737 standard; Protein; 556 AA.
XX AAR96737;
XX AAR96737;
DT 14-AUG-1996 (first entry)
DE XX
XX A. niger Bo-1 carboxypeptidase Y.
XX

Mon Nov 24 13:41:02 2003

KW Carboxypeptidase Y; CPY; ascomycete; deuteromycete; host cell;
 protease deficiency.

XX Aspergillus niger strain Bo-1.

XX WO9609397-A1.

XX 28-MAR-1996.

XX 19-SEP-1995; 95WO-US11945.

XX 20-SEP-1994; 94US-0309341.

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX Thompson SA, Yaver DS;

XX WPI; 1996-108458/19.

XX N-PSDB; AAT28283.

XX Nucleic acid construct encoding a filamentous ascomycete or
 deuteromycete carboxypeptidase Y - useful to produce host cells
 modified to produce reduced amounts of carboxypeptidase

XX Claim 4; Fig 1A-E; 46pp; English.

XX Carboxypeptidase Y (AAR96737), a vacuolar protease, is the product of
 the CPY gene (AAT28283) of *Aspergillus niger* strain Bo-1. Creation
 of CPY-deficient *Aspergillus* strains, e.g. by cloning a selectable
 marker into the CPY gene, provides suitable host strains for prodn.
 of heterologous protein.

XX SQ Sequence 556 AA;

Alignment Scores:

Pred. No.: 7.22e-48 Length: 556
 Score: 630.00 Matches: 154
 Percent Similarity: 48.45% Conservative: 81
 Best Local Similarity: 31.75% Mismatches: 174
 Query Match: 20.81% Indels: 76
 DB: 17 Gaps: 17

US-09-712-338-1 (1-1668) x AAR96737 (1-556)

QY 133 GCAACAAATGTCACCATCCGGTACAGGAACCCGGGCGAGAGCGGTCTGCGAGCTACC 192
 DB 131 AlatyAspLeuArgValLysLysThrAspProGlySerLeuGlyIle-----Asp 147
 QY 193 CCGGTGTCAATCCTACTCTGGATATGTCACACCTCTCC-----GAGTCCCATACCTTC 249
 DB 148 ProGlyValLysGlnTyrThrGlyTyrLeuAspAspAsnGluAsnAspLysHisLeuPhe 167
 QY 250 TTCGGTCTTCGAGCCAGACATACCCAGAACTGCACCTATCACATTTGGTGAAT 309
 DB 168 TyrTrpPhePheGluSerArgAspProGluAsnAspProValValLeuTrpLeuAsn 187
 QY 310 GGTGGCCCTGGAAGCGATCTTTGATCGGTCTCTCGAAGAGTTGGGCCCTTGGCATGTC 369
 DB 188 GlyGlyProGlyCysSerSerLeuThrGlyLeuPheMetGluLeuGlyProSerSerile 207
 QY 370 AATTCGACTTTTGATGACTACATCACCTCCTGCGAAGAGGTCTCCCAATTTACTA 429
 DB 208 AsnLysLysIleGlnProValTyrAsnAspTyrAlaTrpAsnSerAsnAlaSerValIle 227
 QY 430 TTCCTGTCCAGCCATCGGAGTCCGCTTTTCATATAGTATGATGATGGTGGTCCATT 489
 DB 228 PheLeuAspGlnProValAsnValGlyTyrSerTyrSerAsnSerAla-----243
 QY 490 AACCCCTGTAACCTGGGGTCTCGAAAAATTCGAGCTTTGCGAGGAGTTTCAGGGCGGTACCCA 549
 DB 243 -----243
 QY 550 ACCATTGATGCCACTGTGATCGACTACCAATCTTGGCGAGAGCGGCTTGGGAGATC 609

Db 244 -----ValSerAspThrValAlaAlaGlyLysAspValTyrAlaLeu 257
 QY 610 CTGCAAGGATTCCTTAGTGGACTACTAGCTGGACTCTAGGGTGCAGTCTAGGACTTC 569
 Db 258 LeuThrLeuPhePheLysGlnPhePro-----GluTyrAlaLysGlnAspPhe 273
 QY 670 AGTCTATGACGAGAGCTATGAGGGCACTATGGCTCTGCATTTCTCAATCATTTTAC 729
 Db 274 HisIleAlaGlyGluSerTyrAlaGlyHisTyrIleProValPheAlaSerGluIleLeu 293
 QY 730 GAGCAGATGAGAGAAATTCGCAACGGTAGTGTAAATGGTGTTCAGCTTAATTTCAACTCT 789
 Db 294 SerHisLysLysArg-----AsnIleAsnLeuGlnSer 304
 QY 790 CTGGGAATTTAAAGGCATCATCGACGAGGCGATCCAGGCCCTTACTACCTCGAATTC 849
 Db 305 ValLeuIleGlyAsnGlyLeuThrAspGlyTyrThrGlnTyrGluTyrTyrArgProMet 324
 QY 850 GCTGTGAACAATACCTACGGTATCAAGGCTGTCAACGAGACCGTCTACAACTACATGAAG 909
 Db 325 AlaCysGlyAspGlyGlyTyrProAlaValLeuAspLysSerSerCysGlnSerMetAsp 344
 QY 910 TTTCGCAACCAATGCTGCGAGGATTTGATTTCCACCTTCCAAACAGACAAAC 969
 Db 345 -----AsnAlaLeuProArg---CysGlnSerMetIleGluSerCysTyrSerSerGlu 361
 QY 970 CGCACCGCATTTAGCTAGCTAGCCCTCTGCGCGAGACCCACCAACATGTCAGGGACAAT 1029
 Db 362 SerAla-----TrpValCysValProAlaSerIleTyrCysAsnAsnAla 376
 QY 1030 GTTGAGGGGCGCATACTACGCCCTTGTGGTGGTGTGTATGATATTCGGCATCCATAT 1089
 Db 377 LeuLeuAlaProTyrGlnArg---ThrGlyGlnAsnValTyrAspValArgGlyLysCys 395
 QY 1090 GATGACCCG-----ACTCCGCCAAGTTATTACAAATAATTTCTGGCAAG 1134
 Db 396 GluAspSerSerAsnLeuCysTyrSerAlaMetGlyTyrValSerAspTyrLeuAsnLys 415
 QY 1135 GACTGTGTATGAGACGTATCGGGGTCAACATCAC---TACACCCAGTCCATAATGAC 1191
 Db 416 ProGluValIleGluAlaValGlyAlaGluValAsnGlyTyrAspSerCysAsnAsp 435
 QY 1192 GTCTACTACGCTTCCAGCAACAGCGAGCTTTGTCTGGCC-----AACTTCATCGAA 1245
 Db 436 IleAsnArgAsnPheLeuPheHisGlyAspIrpMetLysProTyrHisArgLeuValPro 455
 QY 1246 GACCTCGAGGAGATCTTGTCTCTCCCGTGGGTGTCTCCCTCACTAT---GGCGAGCC 1302
 Db 456 GlyLeuLeuGlu-----GlnIleProVal-----LeuIleTyrAlaGlyAspAla 470
 QY 1303 GATTACATCTGCAACTGTTGCGGGGTTCAGCGGCTTCCCTCGCTGCGAACTACTCCCA 1362
 Db 471 AspPheIleCysAsnTrpLeuGlyAsnLysAlaTrpThrGluAlaLeuGluTrpProGly 490
 QY 1363 GCCGCCAGTTCGCAAGCGAGGTTACAGCGCCCTGAAAGTC-----AAC 1407
 Db 491 GlnAlaGluTyrAlaSerAlaGluLeuGluAspLeuValIleValAspAsnGluHisThr 510
 QY 1408 GCGCTCGAGTATGGGAAACTCGCGAGTATCGTAAATTTCTCTTCCTTCACCTCGCGTCTATGAG 1467
 Db 511 GlyLysLysIleGlyGlnValLysSerHisGlyAsnPheThrPheMetArgLeuTyrGly 530
 QY 1468 GCAGGCCATGAGTCCCATCTACTACAGCCCATCCCTCCCTGCAATGTTTAAACCGGACT 1527
 Db 531 GlyGlyHisMetValProMetAspGlnProGluSerLeuGluPhePheAsnArgTrp 550
 QY 1528 ATCTTCGGT---TGG 1539
 Db 551 LeuGlyGlyGluTrp 555

RESULT 9
 AAR96738

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0126234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148119.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151103.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.


```

Db 51 eCysAspAlaArgSer-----ProGlnPheThrGlyTyrLeuAspIleGlyProLys-- 68
QY 240 CCATACCTCTCTCTGTTCTTCGAAGCCAGACATAACCCAGAACTGCACCTATCACATT 299
Db 69 -HisLeuPhePheIrrPyrPheGluSerGlnAspProPheHisAspProLeuThrLe 88
QY 300 GTGGTTGAAGTGGCCCTCGAAGCGATTCTTTGATCGGTCTCTTCGAAGAGTTGGGCC 359
Db 88 utripMetThrGlyGlyProGlyAspSerSerMetIleGlyLeuPheGluValGlyPr 108
QY 360 TTGCATGTCAT-----TCGACTTTTGATGACTACATCAACCTCTACTCGTGAAGA 413
Db 108 cCysArgIleAsnGluPheGlyAsnGlyThrAspHis---AsnProTrpAlaTrpThrLy 127
QY 414 GGTCTCCAATTACTATTCCNGTCCCGAGCTCGGAGTCCGCTTTCATATAGTGATAC 473
Db 127 sAsnSerSerLeuLeuPheValAspGlnProValAspValGlyPheSerTyrIleAsp-- 146
QY 474 GGTGATGGTGCATTAAACCCGTGAACCTGGGGTCTCGAAAAATTCGAGGAGT 533
Db 147 -----GluGlyTy 149
QY 534 TCAGGGCGGTACCCAAACCATTTGATGCCACTCTGATCATCTACCAATCTTCGCGAGA 593
Db 149 rGlu-----LeuProHisAspSerArgGluAlaAlaVa 160
QY 594 GGCCGCTTGGGAGATCCTCAAGGATTCCTTAGTGGACTACCTAGCTTGGACTTAGGGT 653
Db 160 lAspMetHisArgPheLeuArgLeuPheIleSer-----GluIlePheProHisLy 177
QY 654 GCAGTCTAAGGACTTCAGTCTATGACGGAGAGCTATGGAGGCACTATGGTCTCTGCATT 713
Db 177 sGlnPheLeuProValHisLeuSerGlyGluSerTyrAlaGlyArgTyrIleProTyrLe 197
QY 714 CTTCAATCATTTTACGNGCAGATGAG----- 741
Db 197 uAlaThrGlnIleLeuGluGlnAsnGluLeuTyrLysAspSerProArgIleProLeuLy 217
QY 742 -----AGAATTGCCAACGGTAGTTTAATGGTCTTCAGCTTAATTTCAACTCTCTGGG 794
Db 217 sSerCysLeuValGlyAsnGlyPheMetSerProLysAspAlaThrPheGly-TyrTrpG 237
QY 795 AATTATTAAACGGCATCATCGACGAGGATCCAGGCCCTTACTACCTGAAATTCGCTGT 854
Db 237 luThrLeuCysThrThrAsnSer-GlyValProSerProIle-Phe----- 251
QY 855 GAACAATACCTACGGTATCAAGGTGTCTCAACGAGACCGTCTACAACTACATGAAGTTGC 914
Db 252 -----AsnGluThrArgCysAspIleMet-----Al 260
QY 915 CAACCAATGCCAATGTTCCAGGATTTGATTTCCACCTGCAACACAGACAAACCCGCAC 974
Db 260 aAlaAsnMetProHis---CysMetAspLeuTyrAspIleCysIleGlnHis----- 276
QY 975 CGCATTAGCTACGCGCTCTCGCGCAAGCCACCAACATGTGCGAGGACAATGTGA 1034
Db 277 -----SerAspProAlaIleCysHisAlaAlaGlnSerValCysTyrAspSerValVa 294
QY 1035 GGGGCCATCTACGCTTTGCTGGTGTGTGTATGATATTCGGCATCCATATGATGA 1094
Db 294 lGlyLeuMetAlaLysLeuLeuLeuArg-----MetThrTh 306
QY 1095 CCCGACTCCGCCA-----AGTTATTACAACA 1121
Db 306 rValThrAlaProCysGluIleAspGluMetCysTyrIleGluAlaAlaLeuIleGluAr 326
QY 1122 ATTTGGCAAGGACTGTCTCATGCGAGCTATCGGCTCAACTCAACTACACCCAGTC 1181
Db 326 gTyrLeuAsnSerProSerValTrpGluAlaLeuSerProGlnGlnValThrGluTy 346
QY 1182 CAAT-----AATGACGTCTACTACGCTTTCAGAAACAGCGGACTTTGTCTGSCC 1232
Db 346 rLysPheValAlaThrSerValIleAspAlaPheAlaGlnSerAlaAsp----- 362

```

```

QY 1233 CAACTTCATCGAACACCTCGAGGAGATC-----CTTGCTCTCCCGCTCGGTGTC 1283
Db 363 -GlyMetValSerSerSerLysGlnIleAlaPheLeuLeuAlaAsnAsnValAspPheLe 382
QY 1284 CCTCATCTATGGCGACGCGGATTACATCTGCAACTGGTTCCGGCTCAGGCCGTTTCCT 1343
Db 382 uAlaTyrGlnGlyAsnLeuAspLeuAlaCysAsnThrAlaGlyAsnLeuArgTrpAlaAs 402
QY 1344 CGCTGCCAACTACTCCCAAGCCGCCAGTTCGGAAGCGCAGGTACACGCCCTCGAAAGT 1403
Db 402 nSerLeuSerTrpLysGlyGlnThrGluPheThrAlaLysProLeuLeuProTrpGluIl 422
QY 1404 CAACGGCGTCTAGTATGGGAA---ACTCGCGAGTATGGTAATTTCTCTTCTCATCTCGGT 1460
Db 422 eGln---ValSerValGlyGluGlyThrAspGluThrSerArgPheAlaPheValThrVa 441
QY 1461 CTATGAGCGCAGGCCAT 1476
Db 441 lAspAsnAlaGlyHis 446

```

Search completed: November 21, 2003, 17:40:00
 Job time : 86 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 17:55:56 ; Search time 39 seconds
(without alignments)

3672.287 Million cell updates/sec

Title: US-09-712-338-2

Perfect score: 2979

Sequence: 1 MRGYEFLSVLPVAASWALP.....HTQSSVPLPTATSMSSVGMA 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp rodent:*
- 12: sp.virus:*
- 13: sp vertebrate:*
- 14: sp.unclassified:*
- 15: sp.rvirus:*
- 16: sp.bacteriaph:*
- 17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2976	99.9	555	3 Q96VZ9	Q96VZ9 aspergillus
2	749	25.1	460	3 Q12569	Q12569 absidia zyc
3	601	20.2	552	3 Q96VC4	Q96VC4 emericella
4	596.5	20.0	541	3 Q94152	Q94152 pichia angu
5	595.5	20.0	537	3 Q14414	Q14414 pichia angu
6	479	16.1	507	10 Q8VWQ0	Q8VWQ0 gossypium h
7	473.5	15.9	501	10 Q9XH61	Q9XH61 matricaria
8	473	15.9	429	10 Q8GVT1	Q8GVT1 oryza sativ
9	466.5	15.7	505	10 Q9FFB0	Q9FFB0 arabidopsis
10	464.5	15.6	508	10 Q8L6A7	Q8L6A7 theobroma c
11	439	14.7	510	10 Q9LXH4	Q9LXH4 arabidopsis
12	438	14.7	510	10 Q93ZC3	Q93ZC3 arabidopsis
13	417.5	14.0	2105	5 Q17679	Q17679 caenorhabdi
14	411.5	13.8	499	10 Q9FMX9	Q9FMX9 arabidopsis
15	385	12.9	490	10 Q9FPY7	Q9FPY7 oryza sativ
16	383	12.9	482	10 Q9SV04	Q9SV04 arabidopsis

17	381	12.8	524	10 Q8GTK2	Q8GTK2 oryza sativ
18	364	12.2	480	10 Q9LEY1	Q9LEY1 arabidopsis
19	361	12.1	482	10 Q9FRJ0	Q9FRJ0 oryza sativ
20	360	12.1	437	10 Q9FWG1	Q9FWG1 oryza sativ
21	359	12.1	487	10 Q9SV03	Q9SV03 arabidopsis
22	357.5	12.0	479	10 Q949Q7	Q949Q7 arabidopsis
23	355.5	11.9	360	10 Q9M450	Q9M450 ciccer ariet
24	353	11.8	470	10 Q22732	Q22732 arabidopsis
25	351	11.8	669	10 Q8LPY6	Q8LPY6 oryza sativ
26	349.5	11.7	474	11 Q9D2D1	Q9D2D1 mus musculu
27	348	11.7	452	10 Q9ZQ00	Q9ZQ00 arabidopsis
28	344	11.5	465	10 Q04084	Q04084 arabidopsis
29	344	11.5	494	10 Q9FEU4	Q9FEU4 pisum sativ
30	344	11.5	501	10 Q9SV02	Q9SV02 arabidopsis
31	344	11.5	512	5 Q76725	Q76725 caenorhabdi
32	342.5	11.5	471	10 Q8L9Y0	Q8L9Y0 arabidopsis
33	342.5	11.5	473	10 Q9SGA9	Q9SGA9 arabidopsis
34	337.5	11.3	425	10 Q65568	Q65568 arabidopsis
35	336	11.3	465	10 Q9M099	Q9M099 arabidopsis
36	335.5	11.3	497	10 Q8L7B2	Q8L7B2 arabidopsis
37	333.5	11.2	510	3 Q60123	Q60123 schizosacch
38	333	11.2	487	10 Q9ZUG3	Q9ZUG3 arabidopsis
39	333	11.2	548	4 Q9BR08	Q9BR08 homo sapien
40	331	11.1	465	10 Q94K84	Q94K84 arabidopsis
41	330.5	11.1	472	10 Q9LSM9	Q9LSM9 arabidopsis
42	330.5	11.1	474	10 Q82228	Q82228 arabidopsis
43	330	11.1	445	5 Q814E3	Q814E3 caenorhabdi
44	330	11.1	1203	5 Q45916	Q45916 caenorhabdi
45	328	11.0	504	10 Q9LSV8	Q9LSV8 arabidopsis

ALIGNMENTS

RESULT 1

Q96VZ9 ID Q96VZ9 PRELIMINARY; PRT; 555 AA.
AC Q96VZ9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Carboxypeptidase S1 precursor (EC 3.4.16.6).
GN CPL.
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TK3;
RA van den Broek P.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF394242; AAK77166.1;
DR InterPro; IPR001563; Serine carbpept.
DR Pfam; PF00450; serine carbpept; 1.
DR PRINTS; PR00724; CRBOXYPASEC.
DR PRODOM; PD001189; Serine carbpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase; Hydrolase; Signal.
FT SIGNAL 1
FT POTENTIAL 18
SQ SEQUENCE 555 AA; 61168 MW; 456B63B0CEB55222 CRC64;

Query Match 99.9%; Score 2976; DB 3; Length 555;
Best Local Similarity 99.8%; Pred. No. 7e-201;
Matches 554; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGYEFLSVLPVAASWALPGSTPASVGRRLPKNPTGVKTLTTANNVTIRYKEPGAEGV 60

Db 1 MRGYEFLSVLPVAASWALPGSTPASVGRRLPKNPTGVKTLTTANNVTIRYKEPGAEGV 60

QY 61 CETTTGVKSYSGYVDTSPESHFTFFWFFFAHNPETAPITLWNGGPGSDSLGLFEELGP 120

Db 61 CETTTGVKSYSGYVDTSPESHFTFFWFFFAHNPETAPITLWNGGPGSDSLGLFEELGP 120

121 CHVNSTFDYINPHSWNEVSNLLFLSQPLGVGFSYSDTVGDSINPVTGWVNSFAGYQG 180
121 CHVNSTFDYINPHSWNEVSNLLFLSQPLGVGFSYSDTVGDSINPVTGWVNSFAGYQG 180
191 RYPTIDATLIDTNLAARAAWEILQGLPSLDSRVQSKDFSLWTSYGGHYGPAFFN 240
181 RYPTIDATLIDTNLAARAAWEILQGLPSLDSRVQSKDFSLWTSYGGHYGPAFFN 240
241 HFEYQNERIANGSVNGVQNFNSLGIINGIDEAIQAPYFEFAVNTYGIKAVNETVYN 300
241 HFEYQNERIANGSVNGVQNFNSLGIINGIDEAIQAPYFEFAVNTYGIKAVNETVYN 300
301 YMKFANQMPNGCODLISTCKQTNRTALADYALCAEATNMCRDNVEGPPYAFAGRGVYDIR 360
301 YMKFANQMPNGCODLISTCKQTNRTALADYALCAEATNMCRDNVEGPPYAFAGRGVYDIR 360
361 HPYDDPTPPSYNKKFLAKDSVMDAIGVNTYQSNNDVYAFQQTGDFWPNFIEDLEE 420
361 HPYDDPTPPSYNKKFLAKDSVMDAIGVNTYQSNNDVYAFQQTGDFWPNFIEDLEE 420
421 LALPVRVSLIYGADYICNWFQGOAVSLAANYSOAAQFRSAGYTPKLVNGVVEYGETREYG 480
421 LALPVRVSLIYGADYICNWFQGOAVSLAANYSOAAQFRSAGYTPKLVNGVVEYGETREYG 480
481 NFSFTRVVEAGHEVPYQPIASLQLFNRTIFGWDIAEGOKKIWPSYKNGTATATHTQSS 540
481 NFSFTRVVEAGHEVPYQPIASLQLFNRTIFGWDIAEGOKKIWPSYKNGTATATHTQSS 540
541 VLPPTATSMSSVGMA 555
541 VLPPTATSMSSVGMA 555

RESULT 2
Q12569 PRELIMINARY; PRT; 460 AA.
ID Q12569
AC Q12569
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Prepro-carboxypeptidase Z (EC 3.4.16.5).
GN SPCZ.
OS Absidia zycharae.
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Absidia.
OC NCBI_TaxID=36079;
OX NCBI_TaxID=36079;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRIC 1199;
RX MEDLINE=95308538; PubMed=7788719;
RA Lee B.R., Takeuchi M., Kobayashi Y.;
RT "Molecular cloning and sequence analysis of the scpZ gene encoding the
serine carboxypeptidase of Absidia zycharae."
RL Curr. Genet. 27:159-165(1995).
DR ENBL; D16519; BAA03966.1; --
DR HSP; P00729; 1YSC.
DR MEROPS; S10.010; --
DR InterPro; IPR001563; Serine carboxpept.
DR InterPro; IPR000379; Ser esters site.
DR Pfam; PF00450; serine carboxpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxpept; 1.
KW Carboxypeptidase; Hydrolase.
FT CHAIN 52 460 MATURE CARBOXYPEPTIDASE Z.
SQ SEQUENCE 460 AA; 50825 MW; CE3A410764548B96 CRC64;
Query Match 25.1%; Score 749; DB 3; Length 460;
Best Local Similarity 35.7%; Pred. No. 1.9e-44;
Matches 179; Conservative 72; Mismatches 188; Indels 62; Gaps 15;
12 LVAASWALPGSTASVGRROLPKNPTGKTLTANNVTIRYKPPGABGVCTTPGVKSYK 71

Db 18 LVCAPTVVOAHPMHVLRRQ-DGNDT-----SSGNTTQTKYTSP---KLCD--PDVKQYS 66
Qy 72 GYVDTSPESHFTFFWFFEARHNPETAPITLWNGPGSDSLGLFEEGLPCHVNSTFDYI 131
Db 67 GYLDAANDEHYFFWFFESKNDPKNDPLTILWNGPGSSLLGLWELGPGCQNGS----A 122
Qy 132 NPHSWNEVSNLLFLSQPLGVGFSYSDTVGDSINPVTGWVNSFAGYQRYPTIDATLID 191
Db 123 NPHSWHSSNNMLFFDQPDGVGFSY-----GKQ-----TVS 152
Qy 192 TTNLAAEAABEILQGLPSLDSRVQSKDFSLWTSYGGHYGPAFFNHFEQNERIAN 251
Db 153 TTEDAARANTFLQAFYETFP-----QYKSLDVHFESEYGGHYIPGFASHVDNKKVQS 208
Qy 252 GSVNGVQNFNSLGIINGIDEAIQAPYFEFAVNTYGIKAVNETVYNMKFANQMPNG 311
Db 209 GEKGVVVVPLKSGVGVNGFIDAVIQKSYPRKMTCHSTYPVAVLSE--ECDKMQQIYEND 265
Qy 312 QODLISTCKQTNRTALADYALCAEATNMCRDNVEGPPYAFAGRGVYDIRHPYDDPTPSY 371
Db 266 CKPAEQCAESD-----EDSDCVANQOC-GQIEG-IYAQSGSYFDIRQQGDD-TPHPF 317
Qy 372 YNKFLAKDSVMDAIGVNTYQSNNDVYAFQQTGDFWPNFIEDLEEILALPVRVSLIY 431
Db 318 VDE-LNKASVIKEVGARGHFSMCSVGTAFATGCGA-RSYPVAVKLLKXGIPVLIYV 375
Qy 432 GDADYICNWFQGOAVSLAANYSOAAQFRSAGYTPKLVNGVVEYGETREYGNFTRVYBAG 491
Db 376 GDADYICNWFQGOAVSLAANYSOAAQFRSAGYTPKLVNGVVEYGETREYGNFTRVYBAG 491
Qy 492 HEVPMYQPIASLQLFNRTIFG 512
Db 436 HEVPMYQPIASLQLFNRTIFG 512

RESULT 3
Q96VC4 PRELIMINARY; PRT; 552 AA.
ID Q96VC4
AC Q96VC4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Carboxypeptidase.
GN CpyA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A26;
RX MEDLINE=21333188; PubMed=11440134;
RA Ohsuni K., Matsuda Y., Nakajima H., Kitamoto K.;
RT "Cloning and characterization of the cpyA gene encoding intracellular
carboxypeptidase from Aspergillus nidulans."
RL Biosci. Biotechnol. Biochem. 65:1175-1180(2001).
DR ENBL; AB051820; BAB56108.1; --
DR MEROPS; S10.001; --
DR InterPro; IPR001563; Serine carboxpept.
DR InterPro; IPR000379; Ser esters site.
DR Pfam; PF00450; serine carboxpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxpept; 1.
DR ProSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
SQ SEQUENCE 552 AA; 62065 MW; F67CF39FDBF7D761 CRC64;
Query Match 20.2%; Score 601; DB 3; Length 552;
Best Local Similarity 31.6%; Pred. No. 6.3e-34;
Matches 154; Conservative 77; Mismatches 175; Indels 82; Gaps 19;
Qy 45 ANNVITRYKPPGABGVCTTPGVKSYGVYDTPG-ESHTFFWFFEARHNPETAPITLWLN 103
Db 127 AYDLRIKKTDPSSLGI---DPPVKQYTYGLDDNDKHLFWFFESNDKNDPVPVWLNL 183

127 ITADLRITDSFSPVHAKKGLIRELNLPKGVNVDHRSVLPKGPVKLVERKFPPTLE- 85
60 VCETTPGVKSY-----SGY--VDTSPESTHFFWFPEARHNPETAPITLWNGPGSDSL 111
86 ----VPGVSEFDLGHAGYKLPNSHDARMEYFFESR-NSKKDPVVIWLTGPGCSSE 140
112 IGLFEEGLPCHVNSTDDYINPHSWNEVSNLLFSLQPLGVGFSYSDDTVGGINFVTGVE 171
141 LALFVENGPFITADNMSLVNWEYGMKASNLVYDQPIGTGFSYS----- 185
172 NSSFAGVQGRYPTDITLIDITNLAAEAAEWILQGLFSLGSLDSRVQSKDFSLWTSY 231
186 --SDQDRIHNEDEVSNDLYD-----FLQAFPAHEPF-----AKNDFEITGESYA 229
232 GHYGPFAFFHFHFEQNERIANGSVNGVQVNFNSLGIINGIIIDAIQAPYYPFAVNTYGI 291
230 GHYIPAFARVHRGNK-----AKEGIIHNLKGLFAIGLGLTPAIQYKAYTDYALDVGVIK 284
292 KAVNETVYNYMKFANQMPGNGQDLISTCKQTNR-TALADYALCAEATNMCRDNVEGPPYA 350
285 KSDHRI-----NKLVPVCEMAIKLCGTGDTTISCMASYFVC-----NNIFNGIMA 329
351 FAG-RGVYDIRHPYDDPTPPSYN--KFLAKDSVMDAIGV-NINTQSNNDVYVAFQOTG 406
330 IAGDTNYVDVRKCKCGSLCYDFSNMESFLNKSVRDALGVNLDVFCVCSPTVYQA----- 384
407 DFVWFNFIEDLEIILALPVRVS-----LIY-GDADYICNWFQGVQAVSLAANYSOAAQFRS 460
385 --MLVDWMENLE--VGIPVILLEDGKLLAVYAGEYDLICNLGLNSRWVHAWMSGQKEFVA 440
461 AGYTPKLVNGVEYGETREYGNFSFRVVEAGHEVYQPIASLQLENRTIFGM--DIAG 518
441 SPEVPFIVDGAEGVLKTHGALGFLKVDHAGVPMQPKAALEMJLKRWTGTLSDASDS 500
519 QXKI 522
501 EKLIV 504

RESULT 7
Q9XH61 PRELIMINARY; PRT; 501 AA.
ID Q9XH61 PRELIMINARY; PRT; 501 AA.
AC Q9XH61; 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Serine carboxypeptidase.
OS Matricaria chamomilla.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Asterales; Asteraceae; Asteroidae;
OC Anthemideae; Matricaria.
OX NCBI_TaxID=98504;
RN [1]
RP SEQUENCE FROM N.A.
RA Konchi C., Yasuda H., Hirata T.;
RT "Isolation of a cDNA encoding for a carboxypeptidase, having leucine
RT zipper structure at the N-terminal region, from the cultured shoot
RT primordia of Matricaria chamomilla."
RL Plant Biotechnol. 16:409-412(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Konchi C., Yasuda H., Hirata T.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF141384; AAD42963.2; -
DR HSSP; P00729; 1VSC.
DR MEROPS; S10.009; -
DR InterPro; IPR001563; Serine_carbpept.
DR InterPro; IPR000379; Ser_estrs.site.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXYPEPTASEC.
DR ProDom; PD001189; Serine_carbpept; 1.

64 FWFPEARHNPETAPITLWNGPGSDSLIGLFEELGPCHVNSTDDYINPHSWNEVSNLL 143
148 YWPEARNNDPVNDPVLWNGPGGSSLTGLMFLGSSASIGPDLKPINNPYSWNSNAIVI 207
144 FLSOPLGVGFSYSDTVGDSINPVTGVNENSSFAGVQGRYPTIDATLIDTNLAAEAAWEI 203
208 FLDQPVNVGFSYS-----SKSVGNTVAAGKDVYAF 237
204 LOGELSGLPISLDSRVQSKDFSLWTSYSGHYGPFAFFHFHFEQNERIANGSVNGVQVNFNS 263
238 LELFYQQPFLH-----LKNDFHAGESYGYHYIPVFAEILTHADR-----SFNLT 284
264 LGIINGIIIDEATQAPYYPFAVNTYGIK-AVNETVYNYMKFANQMPGNGQDLISTCKQT 322
285 VLIINGLTDPLNQYFFYERMASCTDGGVEPTLDESECEGM--LETLPK-CLSLIESCYSS 341
323 NRTALADYALCAEATNMCRDNVEGPPYAFAGGVYDIRHP-----YDDPTPPSYNKFEL 376
342 Q-----SVFSCVPASICYCNAQLGPFQK-TGRNVYDVRKVCBGTLCYXD---MEYIDQYL 392
377 AKDSYMDAIGVIN--YTQSNNDVYVAFQOTGDFVAP--NFTEDLEIILALPVRVSLIY- 431
393 NODFVKVKGAEVDYTESCFDVRNRLFPAGDWMKPYHKNVNLLEQ--GLPV---LIYA 447
432 GDADYICNWFQGVQAVSLAANYSOAAQFRSAG-YTELKVNVEYGETREYGNFSFRVVEA 490
448 GDKDFICNWLGNQANSNLPWSGHDEFESAELYNLTLDGTGVGVKNAGKFTFARMFDG 507
491 GHEVYQPIASLQLENRTIFGWDTAEGQK 521
508 GHNVYDQPESSLAMVNRWIAG-DYSLGTKK 537

RESULT 6
Q8VMQ0 PRELIMINARY; PRT; 507 AA.
ID Q8VMQ0 PRELIMINARY; PRT; 507 AA.
AC Q8VMQ0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative serine carboxypeptidase.
GN SCP.
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX NCBI_TaxID=3635;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Xu-142; TISSUE=Fiber;
RA Li C.-H., Zhu Y.-Q., Meng Y.-L., Wang J.-W., Xu K.-X., Zhang T.-Z.,
RA Chen X.-Y.;
RT "Isolation of Genes Preferentially Expressed in Cotton Fiber Cells by
RT CDNA Array and RT-PCR."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY072822; AAL67992.1; -
DR MEROPS; S10.009; -
DR InterPro; IPR001563; Serine_carbpept.
DR InterPro; IPR000379; Ser_estrs.site.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXYPEPTASEC.
DR ProDom; PD001189; Serine_carbpept; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase.
SQ SEQUENCE 507 AA; 56594 MW; 6A0D283D9C5E7713 CRC64;
Query Match 16.1%; Score 479; DB 10; Length 507;
Best Local Similarity 27.4%; Pred. No. 2.1e-25;
Matches 149; Conservative 90; Mismatches 205; Indels 100; Gaps 21;
13 VAASWALPGSTPASVGRRLPKN---PTGVKTLTTANNVTI-----RYKEPGAEG 59

[illegible]

```

RESULT 10
Q8L6A7 PRELIMINARY; PRT; 508 AA.
ID Q8L6A7
AC Q8L6A7;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Carboxypeptidase type III.
OS CP-III.
GN Theobroma cacao (Cacao) (Cocoa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Malvales; Malvaceae; Byttnerioideae; Theobroma.
OC X
OX NCBI_TaxID=3641;
[1]
SEQUENCE FROM N.A.
RP Tazi H., McCarthy J., Bucheli P., Laloi M.;
RA "Molecular characterisation of a type-III carboxypeptidase (cacao CP-
RT III) from Theobroma cacao seeds.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ313514; CAC86383.1; -.
DR InterPro; IPR001563; Serine carboxypeptidase.
DR InterPro; IPR000379; Ser esterase site.
DR Pfam; PF00450; serine carboxypeptidase.
DR PRINTS; PR00724; CB0XYPTASEC.
DR PRODOM; PD001189; Serine carboxypeptidase.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
SQ SEQUENCE 508 AA; 56521 MW; 68B63707DC686BFE CRC64;

Query Match 15.6%; Score 464.5; DB 10; Length 508;
Best Local Similarity 27.8%; Pred. No. 2.2e-24;
Matches 153; Conservative 92; Mismatches 210; Indels 107; Gaps 232

QY 4 YEF-LSVLPLVAASWA-----LPGSTPASVGRRL-----PKNPTGVKLTITA 45
DB 8 YPFFSVSLFLISISSAASFLDRLRGSSPFSFAKKLIRELNLFPKEENVV---VDG 64
QY 46 NNVTI-----RYKEPGAEGVCITTEGVKS-----YSGY---VDTGSDSHVFFHFE 88

```



```

OX NCBI_TaxID=3702;
RN [1]
RA SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koeseema E., Meyers M.C., Banh J.,
RA Bowser L., Carininci P., Dale J.M., Goldsmith A.D., Hayashiraki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J.J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones ";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY057639; ALU15270.1; -.
DR InterPro; IPR001563; Serine carbpept.
DR InterPro; IPR000379; Ser estrs site.
DR Pfam; PF00450; serine carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carbpept; 1.
DR PROSITE; PS00560; CARBOXYPEPT SER_HIS; 1.
DR PROSITE; PS00131; CARBOXYPEPT SER_SER; 1.
SQ SEQUENCE 510 AA; 56837 MW; 704BC00C4E64BF08 CRC64;

```

Query Match	14.7%;	Score 438;	DB 10;	Length 510;
Best Local Similarity	26.6%;	Pred. No. 1.6e-22;		
Matches 123;	Conservative 80;	Mismatches 190;	Indels 69;	Gaps 14;

QY	64	TPGVKSY---SGV--VDTSPESTFWPFPEARINPETAPITLWLNCGPGSDSLIGUFEEL	118
Db	:	: : : : :	:
Db	93	SPSVQDFGHAGYYKLPENSAARMFYFFPESRTN-KADPVVIWLTGGPGGCSSELAIFYEN	151
QY	119	GPCHVNSTFDDYNPHSWNEVSNLLPSOLPGVGFSYSDTVCGSINPVTVCWNSSFAGV	178
Db	:	: : : : :	:
Db	152	GPFTVSNNSSLSNWFEPGWKASNLIIYDOPVGTFGTYSQDSRLRHDEBGV-----	202
QY	179	QGCRYPTIDATLDTNTLAAEAAWEILQLGSLGPSLDSRVQSKDFSILWTSESYGCHYPAF	238
Db	:	: : : : :	:
Db	203	-----SNDLYDFLOAFFFKEHP-----QFVKNDFTITGESYAGHIIPAL	240
QY	239	FNFHYEQNERIANGSVNGVOLFNPSLGIINGIDEAIQAPYYPEFAVNNTYTGIAKVNETV	298
Db	:	: : : : :	:
Db	241	ASRVRHGNK----NKEGTHINKGAIGNGLTNPEIQGVADYALDNKLITQSDDHNL	295
QY	299	YNYMKFANQPMPGCCDLISCTQTNRITALADYA---LCBAETMCRCDNVGEGPYVAAGR-	354
Db	:	: : : : :	:
Db	296	NRY-----YATCQSKIKECSADGEGDACASSYTVC--NNIFQKIMDIAGNV	340
QY	355	GVYDIRHPYDDTPPSYVN--KFLAKDSYMDALGV-NINYTQSNNDVVYAFQOTGFVWP	411
Db	:	: : : : :	:
Db	341	NIYDYVRKQCEGSLCYDFSWMENFLNQKSVRKALGVGDIEFVSCSTAVYAMOMD----	396
QY	412	NTEI-DLEBILALPRVSIYGDADIYICNWFGQGVSLAANTSYQAQFRSAGYTPLKVNG	470
Db	:	: : : : :	:
Db	397	RNLVEVIPALLQDGIKLLVYAGEYDLICNLWGSKWHEMESGQKEFFVAAATVPFHVGN	456
QY	471	VYEGTREXGNPSFTRVVEAGHEVPIYOPIASLQLPNRTIFG	512
Db	:	: : : : :	:
Db	457	KEAGLMKNYGSITFLUKHDAGHWPMDOPKAALQMLQNMQG	498

RESULT	ID	PRELIMINARY	PRT	2105 AA.
Q17679	Q17679			
AC	Q17679			
DT	01-NOV-1996	(TrEMBLrel. 01, Created)		
DT	01-MAY-1999	(TrEMBLrel. 10, Last sequence update)		
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)		
DE	Y16B4A.2	protein.		
GN	Y16B4A.2			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;			
OC	Rhabditidae; Pelodierinae; Caenorhabditis.			
OC	NCBI TaxID=6239;			

RP SEQUENCE FROM N.A.
RA Thomas K.
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z70203; CAA94110.1; -. JOINED.
DR EMBL; AL023825; CAA94110.1; JOINED.
DR EMBL; Z93339; CAA94110.1; JOINED.
DR EMBL; Z93339; CAB07544.1; -.
DR EMBL; AL023825; CAB07544.1; JOINED.
DR EMBL; Z70203; CAB07544.1; JOINED.
DR EMBL; AL023825; CAA19443.1; -.
DR EMBL; Z93339; CAA19443.1; JOINED.
DR EMBL; Z70203; CAA19443.1; JOINED.
DR HSP; P10619; 11VY.
DR WormPep; Y16B4A.2; CB211374.
DR InterPro; IPR001563; Serine carbpept.
DR Pfam; PF00450; serine carbpept; 4.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR PROSITE; PS001189; Serine carbpept; 6.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 4.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 2.
SQ SEQUENCE 2105 AA; 234335 MW; 139CE6E96D75B3FB CRC64;

Query Match 14.0%; Score 417.5; DB 5; Length 2105;
Best Local Similarity 23.5%; Pred. No. 3.6e-20;
Matches 156; Conservative 91; Mismatches 209; Indels 209; Gaps 28;

QY 2 RYEFSLVPLVAASWA-----LPGSTPASVGRQLPKNPTGVKTLTANN 47
Db 1467 QSQVYTTIAGKSWTQNLVLTWKGSHFVSDRPAQ-----ALQMTNFUS 1515
QY 48 VTIRYKEPAEGV-----CET-----TPGV-----KSYGVVDTSPE 80
Db 1516 NOANVSTPAGIDVTPQPLTNVTSQNTCTTGQDTRIINLPLGLPADMLFKCYSGFLDGLSGH 1575
QY 81 HTFFWFEARHNPETAPITLWNGGSDSLGLFEELGFCCHVNSTFDYI-NPHSWNEV 139
Db 1576 KVHYLMVESENPFSTPDLILLWNGGSDSLMGLFEENGPFVRVSKDSQTLSPNYSWNKF 1635
QY 140 SNLFLSQPLGVGFSVDVDSINPTGVNENSSPAGVQGRYPTIDATLDTNLAEEA 199
Db 1636 ANVLYESPIGVGYSA-----YNNTIQYDDVT--TAGE 1668
QY 200 AMEILQGLSLGSLDSRVQSDFLSWTESYGHYGFAPFNHFEYQNERIANGSVNG-VQ 258
Db 1669 NYAALKSFFAQP-----QYTTSDFYTTGESYAGVYLPGL-----SALLVQGIKGDIN 1717
QY 259 LNFNSLGIINGIDE-----ALQAPYPEFAVNNTYGIKAVNETVYN----- 300
Db 1718 INYKVSIGNGVIDKRTDMNSQLHYQY-----HG--GISATTYNTALQLCCSGDEF 1767
QY 301 -----YKFEANQMPNG-----CQD-LISTCKQTKRNTALADYALCAEATNCRDNVEGP 347
Db 1768 KCFSDRMNFNNSIFWGLDSDPCYDFVATGALLTAFDPYNNVQOCWTI----- 1919
QY 348 YFAGAGRVYDIRHPYDDPTPPSYNKK-----FLAKDSV 381
Db 1820 -----PYNDITPRPYGETWTGTYNYESSDFLNGYPCYDDSAMGYLNRPVV 1865
QY 382 MDAIGV--NINTYQSNNDVYACQCGDFWPNFIEDLEILA---LPUEVSLIYGADY 436
Db 1866 RKALNIPDSVPYWAANNIINAYNQVDSIVPN---LQIIMANAPANFKMLLYSGDADT 1921
QY 437 ICNWFQGO-----AVSLAANYSOAA---QFRSAGYTPKLVNGVEYGETREYGNFSFT 485
Db 1922 MNWNLGAEPTANNFAALGTTSSPRAQWTFQYNST-FQP-TVAGYQTSYTNAINIDLVL 1979
QY 486 RYVEAGHEVYYPFIASLQFNRTIFGWDIAGOKKIWPYKXNGTATATHQSSVP--L 543
Db 1980 TVKSGSHFVLDLRPQALQM-----IYNFVKSRGVNTFPDLNSFTTTTTTTTTPGTG 2035
QY 544 PTATS 548

Db 2036 PTVTA 2040

RESULT 14
Q9FMX9 PRELIMINARY; PRT; 499 AA.
AC Q9FMX9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Serine carboxypeptidase II-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98162728; PubMed=9501997;
RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT Sequence features of the regions of 1,191,918 bp covered by seventeen
RT physically assigned P1 clones.";
RL DNA Res. 4:401-414(1997).
DR EMBL; AB007648; BAB11176.1; -.
DR HSP; P08819; 1WHT.
DR MEROPS; S10.005; -.
DR InterPro; IPR001563; Serine carbpept.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00450; serine carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR PROSITE; PS001189; Serine carbpept; 2.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase.
SQ SEQUENCE 499 AA; 56329 MW; 6150FB91A2E7BCBD CRC64;

Query Match 13.8%; Score 411.5; DB 10; Length 499;
Best Local Similarity 26.9%; Pred. No. 1.1e-20;
Matches 152; Conservative 80; Mismatches 216; Indels 117; Gaps 26;

QY 5 EFLSVLPLVAASWALPGSTPASVGRQLPKNPTG-----VKLTANNVTIRKE-PAEG 59
Db 7 EF-SVLVPLVLSFLL-GSTSA-----EKLCSNDGNGCFRSRVLAQAQADRVKELFG--- 57
QY 60 VCETTPGVK--SYSGVVDTSPESH---TFWFEEARHNPETAPITLWNGGSDSL-IG 113
Db 58 ----QPPVFRQVAGIV-TVNETHGALFYWFEEATQNSKPEVLWLNGLNGGPGCSIGFG 112
QY 114 LPEELGFCCHVNSTFDD--YINPHSNVNEVSNLLFLSQPLGVGFSYSTVPGSINPVTGWVE 171
Db 113 AAELGPFPPQNSQPKLKNFYSNMKAANLLFLESFVGFGFSYNT----- 159
QY 172 NSSFAGVQGRYPTIDATLDTNLAEEAAWEILQGLSLGSLDSRVQSDFLSWTESYG 231
Db 160 -----SRDIKQLGDT-VTARDSYNFLVNNFKRFP-----QYKSHDFVIAGESVA 202
QY 232 GHYGFAPFNHFEYQNERIANGSVNQLNFSNGIINGIIDEAIQAPYYPEFAVNNTYGI 291
Db 203 GHVTPQGLSELIYENKIASKDF-----INLKLMIKNALLDDTDQKMEYAWDHAVIS 258
QY 292 KAVNETVYNNYMKFANQM-PNGQODLIS-----TCKQINRTALADYALC 333
Db 259 DALYKVNKNKCDFKQKLVTKECNDALDEYDPVYKILDMYSLYAPKCVPTSTNSSTSHVA 318
QY 334 AEATNCRDNVEGYPYAFAGRVYDIHPYDDPTPPSYNKKFLAKDSVMDAIGNVINTQ 393
Db 319 GNRPLPAFRSILPRILISNEGWRMAAGY-DCASEYTEKYNKDKVQDALHANV----- 373
QY 394 SNNVYVYAFQQTGDFV--W-----PNFTEDLEEIALPVRVSLIYGDADYICNWFQGOAVS 447

205	NK----	DVEKPIILNFKGFMVNAVIDD--	YHDYVGTFEYWWTHGL--	ISD	DTTYQKLQVA	255
306	QY	NQMPN-----	GCQDL--	ISTQKTNR	TALADYALCAEATNMCRDNVGE	357
256	Db	CFESSAHASEACNKIYEVAEAEQGNIDAYS	IVTPTCKKTSFLKRLIRGNSPWLPRGY-	314		
358	QY	DIRHPYDDPTPPSYNKLFLAKDSYMDAIGNIN----	YTSQSNNDVYVAFQQTGFVFWEN	412		
315	Db	-----	DPCTEKYSTKYVNLPEVQKFAHNVGIPYAMTT	CSDDLFUYFKWDSPSRMLPI	367	
413	QY	FIEDLEILALPVRSLIYGDADYICNWFQGGVSL--	AANYSCAAQFRSAGYTPLKNGV	471		
368	Db	Y----	RELIAGLRWVFGSDGDSV-----	VPLTATRYS-----	IDALVLPVTWNY	410
472	QY	EYGETRE-----	YGNFSTRVVEAGHEVPYQPIASQLP	506		
411	Db	PWDDDEVAGWCQVYQGLTIVLTIRGAGHEVFLHPRQALKLF	452			

Search completed: November 21, 2003, 18:01:00
Job time : 41 secs

```

QY 413 FIEDLEEITALPVRVSLIYGDAZYICNWFGGQAVSL-AANYSCAAQFRSAGYTPDKNGV 471
Db 368 Y-----RELIAGLRWVFGDADSV-----VPLTATRYSDALYLPVTWNY 410
QY 472 EYGETRE-----YGNFSFTRVYEAGHEVPPYQPIASLQLF 506
Db 411 PWYDDEEVAGWCQVYQGLTLVTIRGAGHEVELHRPRQALKLF 452

Search completed: November 21, 2003, 18:01:00
Job time : 41 secs

```

This Page Blank (uspto)

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	894	30.0	423	1	CP51_PENJA	P34946	penicillium
2	652.5	21.9	1002	1	CBY9_SCHPO	O13849	schizosacch
3	636	21.3	508	1	YBY9_YEAST	P38109	saccharomyc
4	599	20.1	542	1	CBFY_CANAL	P30574	candida alb
5	594.5	20.0	523	1	CBFY_PICPA	P52710	pichia past
6	527.5	17.7	532	1	CBFY_YEAST	P00729	saccharomyc
7	475	15.9	516	1	CBPX_ARATH	P32826	arabidopsis
8	473	15.9	429	1	CBPX_ORYSA	P52712	oryza sativ
9	455	15.3	500	1	CBPX_ORYSA	P37891	oryza sativ
10	437	14.7	508	1	CBP3_HORVU	P21529	hordeum vul
11	416.5	14.0	500	1	CBP3_WHEAT	P11515	tritium ae
12	359	12.1	476	1	CPVL_HUMAN	Q9h3g5	homo sapien
13	355.5	11.9	470	1	Y5S2_CAEEL	O93991	caenorhabdi
14	353.5	11.9	436	1	CP22_HORVU	P55748	hordeum vul
15	351	11.8	482	1	NF31_NAEFO	P42661	naegleria f
16	349	11.7	474	1	PRTP_MOUSE	P16675	mus musculu
17	347.5	11.7	516	1	CF23_HORVU	P52711	hordeum vul
18	333	11.2	480	1	PRTP_HUMAN	P10619	homo sapien
19	332	11.1	469	1	YUW5_CAEEL	P52717	caenorhabdi
20	331	11.0	729	1	KEX1_YEAST	P09620	saccharomyc
21	326.5	11.0	454	1	YU46_CAEEL	P52715	caenorhabdi
22	318.5	10.7	476	1	CBP2_HORVU	P08818	hordeum vul
23	316	10.6	574	1	YFP3_CAEEL	P52716	caenorhabdi
24	312.5	10.5	471	1	VCP_ADAE	P42660	aedes aegypt
25	312	10.5	510	1	CBP1_ORYSA	P37890	oryza sativ
26	305.5	10.3	423	1	CBP2_WHEAT	O98819	tritium ae
27	297.5	10.0	505	1	YXD2_CAEEL	P52714	caenorhabdi
28	295.5	9.9	286	1	C6PX_PEA	Q41005	pisum sativ
29	280	9.4	499	1	CBP1_HORVU	P07519	hordeum vul
30	271	9.1	523	1	P8FS_ASPPA	P52719	aspergillus
31	255	8.6	452	1	RISC_MOUSE	Q920a5	mus musculu
32	248	8.3	452	1	RISC_RAT	Q920a6	rattus norv
33	237	8.0	452	1	RISC_HUMAN	O9h340	homo sapien

```

178 VQGRYPTIDATLIDTTNLAAEAAWEILQFLSGLPSPDLSRVQSKDFSLWTESYGGHYGPA 237
105 -----TDDVT--STVTAAVYVNNLLQAFYAQRPEYER--DFAIFTESYGGHYGPE 151
238 FENHYFQNERIANGSVNGVOLNFNSLGIINGIIDEAIOAPYPPEFAVNNYIGIKAVNET 297
152 FASYEQOQRAAIKAGSYTGQVNNIVALGVNNGWIDSTIQEKAYIDFSYNNQQIIDGST 211
298 VYNNYMKFANQMPNGQODLISTCKOTNRRTALADYALCAEATNMCRNVEGYPYAFAGRGVY 357
212 RDSLLDAVN---NQCLPALQCCSQSGST-----SDCTWADSVCYQNIEGPISSSGDFVY 263
358 DIRHPYDDPTPPSYNNKFLAKDSVMDAIGVNINYTQSNNDVYVAFQQTGDFVWPNFIEDL 417
264 DIREPSNDPYEPKTYSTVLSDFPTVVYKALGARTNYQECPNGPYNNKPEASTGDNP-RGSLSTL 322
418 BEILLAPVRVSLIYGDADYICNWFGGQAVSLAANSQAAPRSAGYPTPLKVNNGVEYGETR 477
323 SSVVQSGGNVLVWAGDAWICNLWNGYEVANADFPFGNARQFSGALDIAPTVVNGVEXGQPK 382
478 EYGNFSTRVRYEAGHEVPYQPIASLQLFNRTIFGMDIABGQKK 521
383 TVDNFSPFKVYVYGAGHEVPYQPDALQAFKQII-----QKK 418

```

RESULT 2

CBPY_SCHPO
ID CBPY_SCHPO STANDARD; PRT; 1002 AA.
DC 013849; 014366;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carboxypeptidase Y precursor [EC 3.4.16.5] (CPY).
GN CPY1 OR PCY1 OR SPAC19G12.10C.
OS Schizosaccharomyces pombe [Fission yeast].
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
NCBI_TAXID=4896;
RN [1]
RF SEQUENCE FROM N.A., SUBUNIT, SUBCELLULAR LOCATION, AND MUTAGENESIS OF
CY5-627
RX MEDLINE=97352672; PubMed=9209031;
RA Morita M., Iwaihara O., Chikawa Y., Ohuchi N., Sakurai J.-I.,
RA Tsuchi T., Iwahara S., Takegawa K.;
RT "Vacuolar protein sorting in fission yeast: cloning, biosynthesis,
RT transport, and processing of carboxypeptidase Y from
RL Schizosaccharomyces pombe.";
RN J. Bacteriol. 179:4179-4189(1997).
RN [2]
RF SEQUENCE FROM N.A.
RP STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Brooks J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gantley S., Goble A., Hanlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., Mclean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odel C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch B.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds R.G., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grynoprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Mostl D., Hilbert H.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Rodriguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Spakovski G.V., Ussery D., Barrell B.G., Nurse P.; "The genome sequence of Schizosaccharomyces pombe"; Nature 415:871-880(2002).

CC -1- FUNCTION: Involved in degradation of small peptides. Digests preferentially peptides containing an aliphatic or hydrophobic residue in P1' position, as well as methionine, leucine or phenylalanine in P1 position of ester substrate.

CC -1- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a broad specificity.

CC -1- SUBUNIT: Heterodimer of two subunits of 32 kDa and 19 kDa derived from the precursor protein and linked by a disulfide bond.

CC -1- SUBCELLULAR LOCATION: Lysosome-like vacuoles.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D86560; BAA25568.1; --
EMBL; Z97209; CAB10121.1; --
PIR; T43236; T43236.
HSSP; P00729; IYSC.
MEROPS; S10.001; --
GeneDB SPombe; SPAC19G12.10C; --
InterPro; IPR000379; Ser esters site.
InterPro; IPR001563; Serine_carbpept.
Pfam; PF00450; serine_carbpept; 1.
PRINTS; PR00724; CRBOXYPTASEC.
ProDom; PD001189; Serine_carbpept; 1.
PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
PROSITE; PS00560; CARBOXYPEPT SER HIS; FALSE NEG.
KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Repeat; Signal.
FT SIGNAL 1 18
FT PROPEP 19 521 POTENTIAL.
FT CHAIN 522 1002 CARBOXYPEPTIDASE Y.
FT DOMAIN 208 341 PRO-RICH.
FT DOMAIN 225 341 9 X 13 AA TANDEM REPEATS OF M-H-H-E-P-G-E-H-M-P-P-P-P.
FT REPEAT 225 237 1-1.
FT REPEAT 238 250 1-2.
FT REPEAT 251 263 1-3.
FT REPEAT 264 276 1-4.
FT REPEAT 277 289 1-5.
FT REPEAT 290 302 1-6.
FT REPEAT 303 315 1-7.
FT REPEAT 316 328 1-8.
FT REPEAT 329 341 1-9.
FT DOMAIN 361 423 7 X 9 AA TANDEM REPEATS OF D-K-E-H-H-K-G-P-K.
FT REPEAT 361 369 2-1.
FT REPEAT 370 378 2-2.
FT REPEAT 379 387 2-3.
FT REPEAT 388 396 2-4.
FT REPEAT 397 405 2-5.
FT REPEAT 406 414 2-6.
FT REPEAT 415 423 2-7 (APPROXIMATE).
FT ACT_SITE 715 715 BY SIMILARITY.
FT ACT_SITE 921 921 BY SIMILARITY.
FT ACT_SITE 978 978 BY SIMILARITY.
FT BINDING 924 924 SUBSTRATE (BY SIMILARITY).
FT BINDING 979 979 SUBSTRATE (BY SIMILARITY).
FT DISULFID 627 880 BY SIMILARITY.
FT DISULFID 776 789 BY SIMILARITY.
FT DISULFID 799 822 BY SIMILARITY.
FT DISULFID 806 815 BY SIMILARITY.

FT	DISULFID	844	851	BY SIMILARITY.	
FT	CARBOHYD	659	659	N-LINKED (GLCNAC...) (POTENTIAL).	
FT	MUTAGEN	627	627	C->T: 36% OF ORIGINAL ACTIVITY.	
SQ	SEQUENCE	1002 AA;	114237 MW;	4A8D81CFDAB2D854	CRC64;

Query Match	21.9%;	Score 652.5;	DB 1;	Length 1002;
Best Local Similarity	33.8%;	Pred. No. 1.1e-37;		
Matches 165;	Conservative 66;	Mismatches 174;	Indels 83;	Gaps 15;

Qy	49	TIRYKEPCAEGVCETTPGVKSYGVSDTPSPSHTFWFEEFARHPNPETAPITLWNGPGS	108
Db	570	TLRVKDSKPSLGIDT--VKQYTGVLVEDDRHLFFWFEEFSRNDPNDPVVILWNGPGC	627
Qy	109	DSLGLFEEGLPCHVN-STEDDYINPHSWNEVSNLLFSLQPLGVGFSYSDTVGDSINPVT	167
Db	628	SSLTGLEWELGPSSINIETLKPENPHSWNSASVIFLDQPIPTGFSNGD--DSVLDTVT	685
Qy	168	GVVENSFAGVQGRYPIDATLIDTNLAARAAWEILQGLFSLGSLDSRVQSKDFSLWT	227
Db	686	-----AGKDVVAFLNLFPAKFP-----QVAHLDFHIAG	713
Qy	228	ESYGHGYCPAFENHFYQNER----IANG-SVNGVQLNFNSLGTINGIIDEAIQAPYPE	282
Db	714	ESVAGHYVLPQFAKEIMEHNGQANFFVASGEMEKQYINLKSVLGNLTDPLVQYTFYGK	773
Qy	283	FAVNNTYV-----IKAVNETVYVMKFAQMPPNGQDLISTCKQINRTALADYAL	332
Db	774	MACESPYPGIMSQEECDRIITGAVDT-----CAKLITGCYQGTFT-----PV	814
Qy	333	CAEATNMCNRDVGPPYAFAGRGVYDIRHPYDDTPPSY-----YNKFLAKDSVMDAIGV	387
Db	815	CIGASLYCNAMITGP-FTKGLNIYDIRRCRQDEHLCTYPETGAIESYLNQEFVQALGV	873
Qy	388	NINVTQSNDDVYAFQOTGDFVWPNFIEDLEEIL--ALPVRVSLIY-GDADYICNWFQGG	444
Db	874	EYDKGCVNTVIGTFLFGDMWRKTFDDVTAILEAGLPV--LIIAGDADYICNYMGNE	930
Qy	445	AVSLAANYQAAQFRSAGYTPLKVNGVEYGETREYGNFSFTRVVEAGHEVYVYQPIASLQ	504
Db	931	AWTDALAWAQREFYEALKPWSPNGKEAGRSKFNFGVLRLYEAGHMVFPNQPEASLE	990
Qy	505	LENRTIPG 512	
Db	991	MLNSWIDG 998	

RESULT 3				
YBY9 YEAST				
ID	YBY9 YEAST	STANDARD;	PRT;	508 AA.
AC	F38109;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Putative serine carboxypeptidase in ESR1-IRAI inter-genic region			
DE	(EC 3.4.16.-).			
GN	YBR139W OR YBR1015.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
NCBI	Taxid:4932;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN=S288C;			
RX	MEDLINE=94378717; PubMed=8091856;			
RA	Becam A.-M., Cullin C., Grzybowski E., Lacroite F., Nasr F.,			
RA	Ozier-Kalogeropoulos O., Palucha A., Slonimski P.P., Zagulski M.,			
RA	Herbert C.J.;			
RT	"The sequence of 29.7 kb from the right arm of chromosome II reveals			
RT	13 complete open reading frames, of which ten correspond to new			
RL	genes.";			
RL	Yeast 10:S1-S11(1994).			
RN	[2]			
RP	DISCUSSION OF SEQUENCE.			

1	20	POTENTIAL.
FT		
1	21	SIGNAL
FT		
21	111	PROPEP
FT		
112	532	CHAIN
FT		
257	257	ACT SITE
FT		
449	449	ACT SITE
FT		
508	508	ACT SITE
FT		
452	452	BINDING
FT		
509	509	BINDING
FT		
167	409	DISULFID
FT		
304	318	DISULFID
FT		
328	351	DISULFID
FT		
335	344	DISULFID
FT		
373	379	DISULFID
FT		
124	124	CARBOHYD
FT		
198	198	CARBOHYD
FT		
279	279	CARBOHYD
FT		
479	479	CARBOHYD
FT		
24	27	SITE
FT		
508	508	MUTAGEN
FT		
260	261	CONFLICT
FT		
389	389	CONFLICT
FT		
529	529	CONFLICT
FT		

[illegible]

Db 326 ERLGLIESCDSQ-----SWSCVPATYICNNAQLAPYQR-TGRNVYDIRKCEGGLNC 379

QY 366 -PTPPSYNKLAKSDVMDAIGVNVNTQS--NNDVYAFQOTGDFVMPNFIEDLEILAL 423

Db 380 YPTLQD-IDDYLNQDYKVAEADHYEVCNFDINNFILFAGDWMP-PHTAVTDLLNQ 437

QY 424 PVRVSLYGDADYICNWFQGVQVSLAANYSOAAQFRS---AGYPLKVGNGVEYGETREYG 480

Db 438 DLPILVYAGDKDICNWLGNKAWTDVLPWKYDEEFAQKVRNWTASITDEVA-GEVKSXK 496

QY 481 NFGFTVRYEAGHEVPYVQPIASQLNRTIFG 512

Db 497 HFTYLVFNGCHVPPDPVNPENALSMVNEWHIG 528

RESULT 7

CBPX ARATH

ID _CBPX ARATH STANDARD; PRT; 516 AA.

AC P32826; Q9CAE5;

DT 01-OCT-1993 (Rel. 27, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE Serine carboxypeptidase precursor (EC 3.4.16.-)

GN AT3G10410 OR F13M14.32

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Bradley D.;

RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=21016720; PubMed=11130713;

RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseld M.,

RA Farmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,

RA Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,

RA De Simone V., Choise N., Artiguenave F., Robert C., Brottier P.,

RA Wincker P., Cattolico L., Weisenbach J., Saurin W., Quettier F.,

RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,

RA Wurmbach E., Drzonek H., Erle H., Jordan N., Bangert S.,

RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,

RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,

RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,

RA Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Climent J.,

RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,

RA Cooke R., Laudie M., Berger-Liauro C., Purnelle B., Masuy D.,

RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,

RA Monfort A., Argitrou A., Flores M., Lignori R., Vitale D.,

RA Manhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,

RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,

RA Rooney T., Rizzo M., Wals A., Utterback T., Fujii C.Y., Shea T.P.,

RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,

RA Pai G., Millicher J., Sellers P., Gill J.E., Feldblyum T.V.,

RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,

RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,

RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,

RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,

RA Nakayama S., Nakazaki M., Shimp S., Takeuchi C., Wada T.,

RA Watanabe A., Yamada M., Yasuda M., Tabata S.;

RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis

thaliana.";

RL Nature 408:820-822 (2000).

RN [3]

RP SEQUENCE OF 252-372 AND 455-516 FROM N.A.

RC STRAIN=cv. Columbia; TISSUE=Seedling;

RA Hofte H.;

RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.

CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a

broad specificity.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL; M81130; AAB04606.1; -

DR EMBL; AC011560; AAG51389.1; -

DR EMBL; Z55955; CAA81115.1; -

DR EMBL; Z56528; CAA81299.1; -

DR HSSP; P00729; IYSC.

DR MEROPS; S10.009; -

DR InterPro; IPR000379; Ser esters site.

DR InterPro; IPR001563; Serine carboxpept.

DR Pfam; PF00450; serine carboxpept; 1.

DR PRINTS; PR00724; CRBOXYPTASEC.

DR ProDom; PD001189; Serine carboxpept; 1.

DR PROSITE; PS00131; CARBOXYPEPT SER SER; 1.

DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.

KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.

FT SIGNAL 1 22

FT PROPEP 23 82

FT CHAIN 83 516

FT ACT_SITE 229 229

FT ACT_SITE 417 417

FT ACT_SITE 474 474

FT BINDING 420 420

FT CARBOHYD 157 157

FT DISULFID 139 379

FT DISULFID 300 322

FT DISULFID 307 315

FT CONFLICT 515 516

FT QM -> ADVTSSPFLAHENKROOIHKQCVSN (IN REF.

FT 1).

SQ SEQUENCE 516 AA; 57301 MW; 177C778DF657A1C1 CRC64;

Query Match 15.9%; Score 475; DB 1; Length 516;

Best Local Similarity 30.1%; Pred No 1e-25;

Matches 139; Conservative 71; Mismatches 172; Indels 80; Gaps 18;

QY 63 TTPGVKVSQGY--VDTSPESHFFWFPEARHNPETAPITLWNGPGSDSLGLPEELGP 120

Db 93 TVDDLGHAGYKLPKRGASMFYFFESR-NKKDAPVWLTGGGCSSELAVFYENG 151

QY 121 CHVNSTEDDYINPHSNWVSNLLFSLQPLGVGFSYSDVDGVSINPVTGVWVNSSPAGVQG 180

Db 152 FKITSNMLAWNEYGWDQVSNLLYVDQVGTGFSYTTDKSDIRHDETVG----- 200

QY 181 RYPTIDATLIDTTNLAAEAANEILLQGLSLPSLSDRSVQSKDFSLWTSYSGHYGPAPFN 240

Db 201 -----SNDLYDFLQAFPAEPKPL-----AKNDFYITGESYAGHYIPAPAS 240

QY 241 HFYQNERIANGSVNGVOLFNSLGIINGIIDEAIQAPYEPFVAVNNTYGIKAVNNTVN 300

Db 241 RVHKGK--AN---EGVHNLKGFAGNGLTDPALQVPAYPDYALE--MGLITQKE----- 289

QY 301 YMKFANQWPNQCODLISTCKQTNETA-LADYALCAEATNMCRDNVEGVYPAAGRGVYDI 359

Db 290 HDRLEKIVPL-CEUSIKLCGTGDTTCLASLVNCSLFGVWVSHAGGVNY-----YDI 341

QY 360 RHP-----YDDPTPPSYNKLAKSDVMDAIGV-NINYTQSNNDVYVAFQOTGDFWPN 412

Db 342 RKKCVGSLCYDF-----SNMEKFLNQSVRKSLGVGDIDFVSCSTSVYQA-----MLVD 390

QY 413 FIEDLEILALPVRS-----LIY-GDADYICNWFQGVQVSLAANYSOAAQFRSAGYTEL 466

Db 391 WMRNLE--VGIPITLEDGILSVYAGEYDILCNWLNRSRWVNAWESGKTNFGAAKEVPF 448

QY 467 KVGVEGETREYVGNFSFTRVYAGHEVPPYQPIASQLNFR 508

Db 449 IVDGKEAGLLKTYEQLSFLKVRDAGHMVPMQPKAALKMLKR 490

RESULT 8
CBPX ORYSA
ID -CBPX ORYSA STANDARD; PRT; 429 AA.
AC P52712;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase-like precursor (EC 3.4.16.-).
GN CBP31.

OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriactoidae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
[1]
RN RP
SEQUENCE FROM N.A.
RC STRAIN=cv. Yukihikari;
RA Washio K.; Ishikawa K.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
broad specificity.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; D1587; BAA04511.1; -.
DR PIR; T03607; T03607.
DR HSSP; P00729; IYSC.
DR MEROPS; S10.009; -.
DR Gramene; P52712; -.
DR InterPro; IPR000379; Ser estrs site.
DR InterPro; IPR001563; Serine carbpept.
DR Pfam; PF00450; serine carbpept; 1.
DR PRINTS; P00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carbpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydroxylase; Carboxypeptidase; Glycoprotein; Signal.
FT SIGNAL 1
FT CHAIN ? 429
FT ACT_SITE 148 148
FT ACT_SITE 336 336
FT ACT_SITE 393 393
FT BINDING 339 339
FT CARBOHYD 76 76
FT CARBOHYD 414 414
FT CARBOHYD 417 417
FT SEQUENCE 429 AA; 4774 MW; 1D5A668544325B31 CRC64;

Query Match 15.9%; Score 473; DB 1; Length 429;
Best Local Similarity 28.6%; Pred. No. 1.1e-25;
Matches 130; Conservative 77; Mismatches 178; Indels 70; Gaps 16;

QY 70 YSGY--VDTSPEHTFWFPEARNP-ETAPITLWNGGPGSLIGLFEELGCHVNST 126
Db 17 HAGYRLPNTHDARLFYFFEFESRSGKGDDEPVWLTGGPGCSSELALFYENGPFPIADN 76
QY 127 FDIYNPHSNWNEVNLFLQPLCVGFSYSDTVGDSINPVTGVNENSSFAGVQGRPTID 186
Db 77 MSLVWDFGWDQESNLIVDQPTGTGFSYNSNPRDTRHDDAG-VNSDLA----- 125
QY 187 ATLIDTINLAARAAWEILQGLSLGSLDSRVQKPSLMTESYGGHYGPAFFNFHFQON 246
Db 126 -----FLQAFTEHPNF---AKNDYITGESYAGHYIYAFASRVYKGN 165

QY 247 ERIANGSVNQCQVNFNSLGIINGIIDEAIOCAPYYPFAVNTYGIKAVNETVYNYMKFAN 306
Db 166 K-----NSEGIHNLKGFALNGLTDPAIQYKATDYSLD-----MGLITKSQFNRI--N 213
QY 307 QMPNGCQDLISTCKQNR-TALADYALCABATNCRDNVEGPIYAFAGRGVYDIRHP--- 362
Db 214 KIVPTCELAIKLCTSGTISCLGAYVC-----NLICSSIE--TIIGKNVYDIRKPCVG 266
QY 363 ---YDDPTPESYNNKFLAKDSVMDAIGV-NINYTQSNNDVYAFQQTGDFWPNFIE-DL 417
Db 267 SLCYD-----LSNMEKFLQLKSVRESLGVGDIQFVSCSPVYQAML-----LDWMENLEVG 318
QY 418 EEIALPVRVSLYGDADYICNWFPGGQAVSLAANYSOAOFRSAGYTPKLVNGVYGETR 477
Db 319 PELLENDIKVLIVAGEYDLICNWLGNRWNMSMEWSKEAFVSSSEEPFTVDGKEAGILK 378
QY 478 EYGNFSFTRVYEGAGHEVYPTIOPASLQLFNRTIFG 512
Db 379 SYGLSFLKVDHAGHMVPMQPKVALEMLMWTSG 413

RESULT 9
CBP3 ORYSA
ID -CBP3 ORYSA STANDARD; PRT; 500 AA.
AC P37891;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase III precursor (EC 3.4.16.5).
GN CBP3.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriactoidae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
[1]
RN RP
SEQUENCE FROM N.A.
RC STRAIN=cv. Yukihikari; TISSUE=Seed;
RX MEDLINE=92329723; PubMed=1627776;
RA Washio K.; Ishikawa K.;
RT "Structure and expression during the germination of rice seeds of the
gene for a carboxypeptidase.";
RL Plant Mol. Biol. 19:631-640(1992).
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
broad specificity.
CC -!- SUBUNIT: Monomer (Probable).
CC -!- INDUCTION: BY GIBBERELIC ACID (GA). INHIBITED BY ABSICISIC ACID
(ABA).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; D10985; BAA01757.1; -.
DR PIR; S22530; S22530.
DR HSSP; P00729; IYSC.
DR MEROPS; S10.009; -.
DR Gramene; P37891; -.
DR InterPro; IPR000379; Ser estrs site.
DR InterPro; IPR001563; Serine carbpept.
DR Pfam; PF00450; serine carbpept; 1.
DR PRINTS; P00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carbpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydroxylase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1
FT CHAIN ? 500

```

FT PROPEP 22 73 BY SIMILARITY.
FT CHAIN 74 484
FT PROPEP 485 500 BY SIMILARITY.
FT ACT SITE 216 216 BY SIMILARITY.
FT ACT SITE 404 404 BY SIMILARITY.
FT ACT SITE 461 461 BY SIMILARITY.
FT BINDING 407 407 SUBSTRATE (BY SIMILARITY).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 500 AA; 55446 MW; AE455E2780147DB8 CRC64;

Query Match 15.3%; Score 455; DB 1; Length 500;
Best Local Similarity 26.4%; Pred. No. 2.5e-24;
Matches 138; Conservative 85; Mismatches 225; Indels 74; Gaps 17;

QY 7 LSVPLVAASWA-----LP--GSTPASVGR-----OLPK--NPTGVKLTANNVIRY 52
Db 10 LLVVLAASACGLRLPRDAKFAAQAERLIRSNLLPKRAGTGGADVPVAPGELLE 69
QY 53 KPGAEGVCETTPGVKSYG--VDTSPESHPTFFFEARNHNPETAPITLWNGPGSDS 110
Db 70 RRVTLPGDVGDLGHAGYRLPNTHDARMEYFLFESRGKKED-PVWTLTGGPGCSS 128
QY 111 LIGLFEELGPGCHVNSTDDVINPHSWNEVSNLLFLSOPLGVGFSYSDTVGGINPTGVV 170
Db 129 ELAVFYENGFTISNNMSLAWNKEGWDITISNIFVDQPTGTGFSYSSDDRDTRHDETV- 187
QY 171 ENSFAGVQGRYPTIDATLIDTTNLAAEAAWEILQGLSLGSLDSRVQSKDFSLWTESY 230
Db 188 -----SNDLYSFLQVFFKKHPEF-----AKNDFFITGESY 217
QY 231 GGHVGAFFNHFYEQNERIANGSVNGVQVNFSLGIINGIIDRAIQAPYYPFAVNNYTG 290
Db 218 AGHYIFAFARVHGQGNK--AN--EGHINLKGFAIGNGLTDAIQYKAYTDVALD--- 268
QY 291 IKAVNETVVMYMKFANQMPNGCODLSTCKQNTALADYALCAEATNMCRDNVEGPYA 350
Db 269 MNLIKKSDYDRI---NKFIPPCFAIKLCTNGK-----ASCWAAVMVNCISFSSIMKL 319
QY 351 FAGRGVYDIRHPYDDPTPPSYN--KFLAKDSYMDAIGV--NINYTQSNNDVYVAFQQTGD 407
Db 320 VGTNTYDVRKECEGKLCYDFSNLEKFFGDKAVKEAIGVGDLFFVCSITVYQAMLTD-- 377
QY 408 FVWPNFIE-DLEILALPVRSLIYGADYICNWFQGVQVSLAANYSOAAQFSAQYTPL 466
Db 378 --WMRNLEVGIPALLEDGINVLVYAGEYDLICNWLGNRSRWHSMEWSGQKDFVSSHESFP 435
QY 467 KVNVEYGETREYGNFSFTRVYEAQHEVPVYQPIASLQLFNR 508
Db 436 VVDGAEGVLKSHGSLFLKVNAGHVMVDQPKASLEMLRR 477

RESULT 10
CBP3 HORVU
ID CBP3 HORVU STANDARD; PRT; 508 AA.
AC P21529;
DT 01-MAY-1991 (Rel. 18, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase III precursor (EC 3.4.16.5) (CP-MIII).
GN CBP3 OR EXP;3.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Himalaya; TISSUE=Aleurone;
RA Rocher A., Lok F., Cameron-Mills V., von Wettstein D.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 81-491.
RC STRAIN=cv. Gula;

```

```

RX MEDLINE=903115015; PubMed=2639682;
RA Soerensen S.B., Svendsen I., Breddam K.;
RT "Primary structure of carboxypeptidase III from malted barley.";
RL Carlsberg Res. Commun. 54:193-202(1989).
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
CC broad specificity.
CC -!- ENZYME REGULATION: INHIBITED BY MERCURIC IONS.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: SECRETED INTO THE ENDOSPERM.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED MAINLY IN THE ALEURONE AND, TO A
CC LESSER EXTENT IN THE EMBRYO, THROUGHOUT THE 5-DAYS GERMINATION
CC PERIOD EXCLUSIVELY WITH A MAXIMAL LEVEL AT 3 DAYS. ALSO FOUND IN
CC THE ROOTS AND SHOOTS OF THE GROWING SEEDLING.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y09604; CAA70817.1; .
DR HSP; P00729; ICFY.
DR MEROPS; S10.009; .
DR InterPro; IPR000379; Ser_estr site.
DR InterPro; IPR001563; Serine carboxpept.
DR Pfam; PF00450; serine carboxpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine carboxpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT SER_SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT SER_HIS; 1.
KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 80
FT CHAIN 81 491 SERINE CARBOXYPEPTIDASE III.
FT PROPEP 492 508
FT MOD_RES 81 81 BLOCKED.
FT ACT SITE 223 223 BY SIMILARITY.
FT ACT SITE 411 411 BY SIMILARITY.
FT ACT SITE 468 468 BY SIMILARITY.
FT BINDING 414 414 SUBSTRATE.
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .).
FT VARIANT 265 265 Q -> V (POLYMORPHISM).
SQ SEQUENCE 508 AA; 56362 MW; 70C6751D78D40AB6 CRC64;

Query Match 14.7%; Score 437; DB 1; Length 508;
Best Local Similarity 26.5%; Pred. No. 4.5e-23;
Matches 134; Conservative 78; Mismatches 225; Indels 68; Gaps 15;

QY 12 LVAASWALPGSTPASVGRRLPKNTGVKLTANNVIRYKEPG-AGEVCETTPGVKSY 70
Db 40 LIRALLPKDSSSSSSGRHG-ARVCEGNEDEVAPGQLLERRVTLPGLPADVAD---LGH 94
QY 71 SCY--VDTSPESHPTFFFEARNHNPETAPITLWNGPGSDSLIGLFEELGPGCHVNSTFD 128
Db 95 AGYRLPNTHDARMEYFFFSRGKKED-PVWTLTGGPGCSSELAFFVENGPFITANNMS 153
QY 129 DYINPHSWNEVSNLLFLSOPLGVGFSYSDTVGGINPTGVVNSVNSFAGVQRYPTIDAT 188
Db 154 LVWNKFGWKDISNIIFVDQPTGTGFSYSSDDRDTRHDETV----- 194
QY 189 LIDTTNLAAEAAWEILQGLSLGSLDSRVQSKDFSLWTESVGHYGAFFNHFYEQNER 248
Db 195 -----SNDLYDFLQVFFKKHPEF-----IKNDFFITGESYAGHYIFAFARVHGQGNK 242
QY 249 IANGSVNGVQVNFSLGIINGIIDRAIQAPYYPFAVNNYTGIKAVNETVVMYMKFANQ 308
Db 243 -----NEGTHINKGFAIGNGLTDAIQYKAYTDVALEMNLIQKADYERI-----NKF 290
QY 309 PNGCQDLISTCKQNTNR-TALADYALCAEATNMCRDNVEGPYVAFAGRGVYDIRHEYDDPT 367

```

Query Match 14.0% Score 416.5; DB 1; Length 500;
Best Local Similarity 24.6%; Pred No. 1.2e-21;
Matches 131; Conservative 85; Mismatches 231; Indels 85; Gaps 16;

QY 7 LSVLPLVAAGWALPGSTPASVGRRLPKNPTG-----VKLTUTANNVTIRYKPGAE 58
Db 1 MATTPRLASLLALLALCAAAGALRLPPDASFPQAQAERLIRALNLLPGRPRGLGAGAE 60
QY 59 GVCE-----TTPGVKS-----YSGY--VDTSPESHPTFFWFFSARHNPTAPITLM 101
Db 61 DVAPQLLERVTLPLGLPEGVGDIHGHEAGYRLNTHDARMFYFFFSRGKKED-PVVIIW 119
QY 102 LGGPGSDSLGLPEELGCHVNSTEDDYINPHSWNEVSNLLFLSQPLGVGSYSDTVDG 161
Db 120 LTGGPGSSSELAVFYENGPTTIANNMSLVNWKFGWDKISNIIFVDPATGTGFSYSSDDRD 179
QY 162 SINPVTGVNENSPAGVQGRYPTIDATLIDTNTLAANAWEILOGLFSLGSLDSRVOSK 221
Db 180 TRHDEAGV-----SNDLYDFLOVFFKKHEF-----VKN 208
QY 222 DFLWTSYGYGHPAPENHFEYQNERIANGSVNGVQLNFSLGLINGIIIDEAIOAPYYP 281
Db 209 DFFITGESYAGHYTPAFASRVHQGKK-----NEETHNLKGFALCNGLTDPALQYKAT 263
QY 282 EFAYNTYGIKAVNETVYNNMKEFANQMPGCGODLISTCKQTNR-TALADYALCAEATNMC 340
Db 264 DYALD-----MNLQKADYDRI---NKFTPPCEFAIKLCGTGKASCMAAYMVCNSIFNSI 316
QY 341 RDNVEGPPYAFAGRGVYDIRHPYDPTPPSYN--KFLAKOSVMDAIGV-NINYTQSNND 397
Db 317 MKLVGTKNY-----YDVRKECEKLCYDFSNLEKFECDKAVROAIGVGDIEFVSCSTS 369
QY 398 VYIAFQQTGDFWENFIB-DLEIILALPVYSLIYGDADYICNWFPGGQAVSLAANYSOAA 456
Db 370 VYQAMLT-----WMENLEVGIPALLEDGINVLIVAYGEYDLICNWLGNRWVHSMWESGQK 425
QY 457 QFRSAGYPLKNGVGEYGETREYGNFSTRVVEAGHEVPYQFIASLQLENR 508
Db 426 DPAKTAESFLVDDAQAGVILKSHGALSFLKVNAGHVMWMDQPKAALMLRR 477

RESULT 12
CPVL_HUMAN STANDARD; PRT; 476 AA.
ID CPVL_HUMAN
AC Q9H3G5; Q8HBL7; Q96AR7; Q9HB41;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable serine carboxypeptidase CPVL precursor (EC 3.4.16.-)
DE (Carboxypeptidase, vitellogenic-like) (Vitellogenic carboxypeptidase-like protein) (VCP-like protein).
DE CPVL OR VLP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21295045; PubMed=11401439;
RA Mahoney J.A., Ntoliou B., Dasilva R.P., Gordon S., McKnight A.J.;
RT "Cloning and characterization of CPVL, a novel serine
RT carboxypeptidase, from human macrophages";
RL Genomics 72:243-251 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX Cho J.-J., Baik H.-H.;
RT "Cloning of VCP-like protein expressed in human heart and placenta";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Placenta;
RA Ota T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayashi K., Ishii S.,
RA Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,

291 IPPCEFAIKLCGTNGKASCMAYMVCNTIFNSIMKLVTGKNY-----YDVRKECEGKL 343
368 PPYYN--KFLAKOSVMDAIGV-NINYTQSNNDVYAFQGTGDFWENFIE-DLEIILAL 423
344 CYDSNLEKFGDKAVRQAIGVGDIEFVSCSTSIVQAVLTD-----WNRNLEVGIPALLED 399
424 PVYSLVIGDADYICNWFPGGQAVSLAANYSOAAQFRSAGYTPLVKNGVGEYGETREYGNFS 483
400 GINVLIVAGEYDLICNWLGNRWVHSMWESGQKDFAKTAESSFLVDDAQAGVILKSHGALS 459
484 FTRYEAGHEVPYQFIASLQLENR 508
460 FLKVNAGHVMWMDQPKAALMLRR 484

RESULT 11
CBP3 WHEAT STANDARD; PRT; 500 AA.
ID CBP3 WHEAT
AC P11515;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase III precursor (EC 3.4.16.5) (CP-WIII).
GN CBP3.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88007602; PubMed=2820978;
RA Goulcombe D.C., Barker R.F., Jarvis M.G.;
RT "A gibberellin responsive wheat gene has homology to yeast
RT carboxypeptidase Y";
RL J. Biol. Chem. 262:13726-13735 (1987).
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
CC broad specificity.
CC -!- SUBUNIT: Monomer (Probable).
CC -!- INDUCTION: By gibberellic acid (GA).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J02817; AAA34273.1; -;
CC FIR; A29412; A29412.
CC HSP; P00729; ICPY.
CC MEROPS; S10.009; -;
CC InterPro; IPR000379; Ser_estrs_site.
CC InterPro; IPR001563; Serine_carbpept.
CC Pfam; PF00450; serine_carbpept; 1.
CC PRINTS; PR00724; CBPOXYPTASEC.
CC PRODOM; PD001189; Serine_carbpept; 1.
CC PROSITE; PS00131; CARBOXYPEPT_SER_HIS; 1.
CC PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
CC K0 Hydroxylase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 21
FT PROPEP 22 73 BY SIMILARITY.
FT CHAIN 74 484 SERINE CARBOXYPEPTIDASE III.
FT PROPEP 485 500 BY SIMILARITY.
FT ACT SITE 216 216 BY SIMILARITY.
FT ACT SITE 404 404 BY SIMILARITY.
FT ACT SITE 461 461 BY SIMILARITY.
FT BINDING 407 407 SUBSTRATE (BY SIMILARITY).
FT CARBOHYD 144 144 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 500 AA; 55334 MW; B2ACE10EF8484CDA CRC64;

RA Nagahari K., Sugano S., Isogai T.;
 RT "HRI human cDNA sequencing project."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abrancini P., Prange C.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: May be involved in the digestion of phagocytosed
 CC particles in the lysosome, participation in an inflammatory
 CC protease cascade, and trimming of peptides for antigen
 CC presentation.
 CC TISSUE SPECIFICITY: Expressed in macrophages but not in other
 CC leukocytes. Abundantly expressed in heart and kidney. Also
 CC expressed in spleen, leukocytes, and placenta.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF106704; AAC37991.1; -;
 DR EMBL; AF282617; AAG14348.1; -;
 DR EMBL; AK075433; BAC11618.1; -;
 DR EMBL; BC016838; AAH16838.1; -;
 DR HSSP; P10619; IIVY.
 DR MEROPS; S10.003; -;
 DR Genew; HGNC:14399; CPVL.
 DR InterPro; IPR000379; Ser_estr site.
 DR InterPro; IPR001563; Serine_carbpept.
 DR Pfam; PF00450; serine_carbpept; 1.
 DR PRINTS; PR00724; CRBOX1PTASEC.
 DR Prodom; PD001189; Serine_carbpept; 1.
 DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; FALSE_NEG.
 KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
 FT SIGNAL 1 22
 FT PROPEP 23 ? POTENTIAL.
 FT CHAIN ? 476 PROBABLE SERINE CARBOXYPEPTIDASE CPVL.
 FT ACT_SITE 204 204 BY SIMILARITY.
 FT ACT_SITE 388 388 BY SIMILARITY.
 FT ACT_SITE 448 448 BY SIMILARITY.
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 346 346 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 25 25 R -> H (IN REF. 2).
 FT CONFLICT 284 284 L -> F (IN REF. 3 AND 4).
 FT CONFLICT 287 287 F -> L (IN REF. 3).
 FT CONFLICT 398 398 H -> R (IN REF. 3 AND 4).

CC	EMBL: U28730; AAA68259.1; --	
DR	PIR: T16606; T16606.	
DR	HSP; F10619; LIVY.	
DR	MEROPS; S10.002; --	
DR	WormPep; K10B2.2; CE02009.	
DR	InterPro; IPR00379; Ser estrs site.	
DR	InterPro; IPR001563; Serine carboxypept.	
DR	Pfam; PF00450; serine carboxypept; 1.	
DR	PRINTS; PR00724; CRBOXPTASEC.	
DR	ProDom; PD001189; Serine carboxypept; 2.	
DR	PROSITE; PS00131; CARBOXYPEPT SER; 1.	
DR	PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.	
KW	Hypothetical protein; Hydrolase; Carboxypeptidase; Glycoprotein;	
KW	Signal.	
FT	SIGNAL	1 19
FT	CHAIN	20 470
FT	ACT SITE	169 169
FT	ACT SITE	380 380
FT	ACT SITE	441 441
FT	ACT SITE	132 132
FT	CARBOHYD	316 316
FT	CARBOHYD	396 396
FT	CARBOHYD	470 470
SQ	SEQUENCE	470 AA; 53158 MW; CCC2DACB75EF30FC CRC64;
	Query Match	11.9%; Score 355.5; DB 1; Length 470;
	Best Local Similarity	25.2%; Pred. No. 1.9e-17;
	Matches 140; Conservative	69; Mismatches 194; Indels 153; Gaps 22;
QY	5 EFLSVLPVLAASWALPGTSPASVGRRLPKNPTGVKTLTANNVIRIKPEAGVCHT	64
Db	2 KLLSLIFVSSYFCLAAPATDKVNDLP	35
QY	65 PVKSYGVDTSPESHFFFEARHNPETAPITLWNGPGSDSLIGLPEELGFCVHN	124
Db	36 PDFHYGYLRAWTDKLYHWLTSSRAPQDPLVWLNNGPGCCSLDLIELGPFHVK	95
QY	125 STFDD-YINPHSWNEVSNLLFLSQPLGVGFSYSDTVGDSINPVTGVVENSFAGVGRYP	183
Db	96 DFGNSIYYEYAWNKEFANVLFLESFAGVGYSTNPNLTVSDDEVSLHN	144
QY	184 TIDATLIDTNLAEEAWELIQFLSGLPSLSRVQSKDFSLWTSYGHYGPAPFNHY	243
Db	145 -----YMALDFLSKFPYKGR-----DFWITGESYAGVYPTL-----	178
QY	244 EQNERIANGSVNGVQVNLNENSLGII-----DEALQAPYPEFAVNNYGIKAVNETVY	299
Db	179 --AVRLNDKKNP--NFKGVALGNALPNPNNTWVPY-----YYHALVRDDLY	226
QY	300 NYMKFANQPNQGDLLISTCKQTNRTALADYALCABATNCRDNVEGYPYAFAGRVYDI	359
Db	227 N-----DIARNCCNNIGTCD-----IYSKFPDPN--CRDKV--INALDGTNELNM	268
QY	360 RHYDDPTPPSYN-----KFLAKOSVMDAIGV-----NINVT-----QSNNDVYV	400
Db	269 YNLYD-----VCYYNPTNLKAFIERQMRIAVALGPKARKNAATTVFLCAQTNNTHVYLN	324
QY	401 -----AFQQTGDFVWPNF-----IBDLREILALPVRVSLIYGDADYICN	439
Db	325 ADVKRSKLHLPSSLPAMEECSQVGKYNVTHFNVIFEFQTMIAAGIKILVYNGVDVTACN	384
QY	440 WFGQGVASLAANYSSQAQ-----FRSAGYPLKNGVVEYGTREYGNPSFTRVVEAGH	492
Db	385 SIMNQFLTSLNLTVLGEQEKVNEAWHYSQGTGTAVAGF---QTKFAGNVDFLTVRGS	441
QY	493 EVYYQBIASLQ-LFN 507	
Db	442 FVPEDKPKESQOMIFN 457	
	RESULT 14	
	CF22_HORVU	
ID	CF22_HORVU	
	STANDARD; PRT; 436 AA.	

AC	P55748;	
DT	01-NOV-1997 (Rel. 35, Created)	
DT	01-NOV-1997 (Rel. 35, Last sequence update)	
DT	28-FEB-2003 (Rel. 41, Last annotation update)	
DE	Serine carboxypeptidase II-2 precursor (EC 3.4.16.6) (CP-MII.2)	
DE	(fragment).	
DE	CXP;2-2.	
GN	Hordeum vulgare (Barley).	
OS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;	
OC	Triticeae; Hordeum.	
OX	NCBI_TaxID=4513;	
RN	[1]	
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.	
RC	STRAIN=cv. Alexis; TISSUE=Grain;	
RC	MEDLINE=94336715; PubMed=7520177;	
RA	Dal Degán F., Rocher A., Cameron-Mills V., von Wettstein D.;	
RT	"The expression of serine carboxypeptidases during maturation and	
RL	germination of the barley grain."	
CC	Proc. Natl. Acad. Sci. U.S.A. 91:8209-8213(1994).	
CC	-!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine	
CC	or lysine residue.	
CC	-!- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS	
CC	COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND (BY SIMILARITY).	
CC	-!- DEVELOPMENTAL STAGE: EXPRESSED IN THE GERMINATING EMBRYO. LOW	
CC	LEVELS IN THE DEVELOPING ALEURONE AND EMBRYO. ALSO FOUND IN THE	
CC	ROOTS AND SHOOTS OF THE GROWING SEEDLING.	
CC	-!- PTM: THE LINKER PEPTIDE IS ENDOPROTEOLYTICALLY EXCISED DURING	
CC	ENZYME MATURATION (BY SIMILARITY).	
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; X78878; CAB59202.1; --	
DR	HSP; P08819; 1WHT.	
DR	InterPro; IPR00379; Ser estrs site.	
DR	InterPro; IPR001563; Serine carboxypept.	
DR	Pfam; PF00450; serine carboxypept; 1.	
DR	PRINTS; PR00724; CRBOXPTASEC.	
DR	ProDom; PD001189; Serine carboxypept; 1.	
DR	PROSITE; PS00131; CARBOXYPEPT SER; 1.	
DR	PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.	
KW	Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Multigene family.	
FT	NON_TER	1 256
FT	CHAIN	<1 256
FT	PROPEP	257 270
FT	CHAIN	271 436
FT	ACT SITE	149 149
FT	ACT SITE	350 350
FT	ACT SITE	403 403
FT	DISULFID	56 333
FT	DISULFID	217 229
FT	DISULFID	253 281
FT	CARBOHYD	107 107
SQ	SEQUENCE	436 AA; 48952 MW; E0F82D97E0C34DC9 CRC64;
	Query Match	11.9%; Score 353.5; DB 1; Length 436;
	Best Local Similarity	25.7%; Pred. No. 2.4e-17;
	Matches 123; Conservative	76; Mismatches 173; Indels 107; Gaps 23;
QY	70 YSGYVDTSP--SHTFFWFEEARHNPETAPITLWNGPGSDSL-IGLPEELGFCVHNST	126
Db	16 YAGYVTVSEDRGALFWFEAAHDPASKPELLWNGPGCSIAFGVGEVGFPHVAD	75
QY	127 FDD-YINPHSWNEVSNLLFLSQPLGVGFSYSDTVGDSINPVTGVVENSFAGVGRPTI	185
Db	76 GKGVHNPYSNWNOVANILFLDSEVGVGYSYNT-----SADILSN-----GDERTA	121

QY	186	DATLDTNTLAABAWEILLQGLSLGSLDSRVQSKDFSLWTSETSYGGHYGPAFFNHFYEQ	245
Db	122	KDSLFLTK-----W--LRF-----	
QY	246	NERIANGSVNGVQNFNSGIINGIIDEIAIQPYTPEFANNYTGKAVNTEVYNYMKFA	305
Db	166	HEATGKSI-----NLKGYMVGNAITDDP--HDHYGIFYQYMWTTGL--ISDQTYKLLNIF	216
QY	306	NQMPNGC--QDLISCTQNRUALADYALCAEATNCRDNVZGP--YYAFAG-----	353
Db	217	-----CDFSPVHTSPQCDK--ILDIA--STEAGNIDSYSIEFTPTCHSFASSRNKVKVR	267
QY	354	RGVVDIRHPYDDPTPPS--YYNFKLAKDS--VMDAIG-----VNINYTQSND	397
Db	268	LRSGVKMGCEQDPCTEKHSIVVFNLHEVQKALHVPVIGKSKWETCSEVINTNWKDCERS	327
QY	398	VYAFQQTGDFWPNFIEDLEBILAPRVSLIYGDADYICNWFQGVQASLAANTYSQAAQ	457
Db	328	VLHIY-----HELIQYGRIMVFSGTDVAV-----IPVTSRYSIDA	364
QY	458	FRSAGYTPLKV---NGVEYGETREYGVNFSFTRVYEAGHEVPYQIOPIASLOFNRTIFG	512
Db	365	LKLPVTFEWHAYDDGEGVGWTCQYKGLNFVTVRGACHEVPLRPKQALLTIKSLFLAG	423

RESULT 15

NS011_13					
NF31_NAEFO					
ID	NF31_NAEFO	STANDARD;	PRT;	482 AA.	
AC	A42661;				
DT	01-NOV-1995 (Rel. 32, Created)				
DT	01-NOV-1995 (Rel. 32, Last sequence update)				
DE	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Virulence-related protein NF314 (EC 3.4.16.-).				
OS	Naegleria fowleri.				
OC	Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.				
OX	NCBI_TaxID=5763;				
RN	[1]				
RP	SEQUENCE FROM N.A.				

```

SEQUENCE FROM N.A.
RC STRAIN=LEE;
RA MEDLINE=92267659; PubMed=1587609;
RA Hu W.-N., Kopachik W., Band R.N.;
RT "Cloning and characterization of transcripts showing
RT virulence-related gene expression in Naegleria fowleri.";
RL Infect. Immun. 60:2418-2424(1992).
CC
CC -1- FUNCTION: MAY BE REQUIRED BUT IS NOT SUFFICIENT FOR INCREASED
CC VIRULENCE.
CC
CC -1- INDUCTION: BY GROWTH ON MAMMALIAN CELLS.
CC
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M88397; AAA29384.1; -.
DR PIR; A43828; A43828.
DR HSP; P10619; LIVV.
DR MEROPS; S10.UPW; -.
DR InterPro; IPR000379; Ser estrs site.
DR InterPro; IPR001563; Serine_carbpept.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXYPASEC.
DR ProDom; PD001189; Serine carbpept; 2.
DR PROSITE; PS001131; CARBOXYPEPT SER SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
DR Hydrolase; Carboxypeptidase.
DR ACT_SITE 163 163 BY SIMILARITY.
DR ACT_SITE 399 399 BY SIMILARITY.
DR ACT_SITE 459 459 BY SIMILARITY.

```

SQ	SEQUENCE	482 AA;	53848 MW;	0B83049C1A8A0908 CRC64;
Query Match	11.8%;	Score 351;	DB 1;	Length 482;
Best local Similarity	25.0%;	Pred. No. 4e-17;		
Matches	139;	Conservative	57;	Mismatches 165; Indels 196; Gaps 23;
QY	55	PGAGGVCETTPGVKSGSYV--DTSPESTTFFWFFFEARNNPETAPITLWLNGPFGSDSLI	112	
Db	20	PGLSG---NIGVKSYGVLANATGRFLFYWFEEFSMRNPSODPLVMVTGGPFGCSLIG	75	
QY	113	GLFEEELGPCHVNGSTEDDYI-NPHSWEVSNLLFLSQPLGVGSYSQTDVGDSINPVTCGVE	171	
Db	76	GEASEHGLFLVNADGATITRNPSWNRVSNILTYEQPVGVGSYSNST---	124	
QY	172	NSSFAGVQGRYPIDATLIDTTNL-----AAEAWEILOGLSLGPLSDSRVQSOKDFSLWT	227	
Db	125	-----DYQNLDNDVQAASDMNNALRDLTRFPQFIGR-----ETYLAG	161	
QY	228	ESYGCHGP-APFNHFYEQNERIANGSVNGVQLNFNSLGII--NGIID-----	272	
Db	162	ESYGCVYVTTAVN-----IVEGNKGQQPYVNLVGLIVNGVGTDAEADSINSIFPMW	213	
QY	273	---EAIQAPYPPEFAVMNTYGIKAVNETVYNMKAFANQMPNGCODLIISTCKQTNRITALD	329	
Db	214	KYHLSISIKYYEB-----GYKACQGDFY----ANQNLPAQCREFL-----	249	
QY	330	YALCAEATNWCIRDNVGPPYAF-----	351	
Db	250	-----DSSN-AMGNI-NPYIYDSCFWLGINLQQLKTQEMTFQVLDPKTOPVKIHPL	302	
QY	352	-----AGRGVDIRHPDDPTPPSY-YNKLEKADSVMDALGVN-----	388	
Db	303	FQMVKHGWSKRVANERNAFAPFETDAPCVNQSTAKYFRRLDVQOALGVRRKRTADPNGW	362	
QY	389	-----INTQSNNDV--YYAFQQTGDFFWMFNFIEDLEEIALP-VRYSLIYGDADIYCN	439	
Db	363	NICTGHINYQTVYSTILPFYA-----KLLPHIRILVYSGDITDMVVN	403	
QY	440	WFGQAVSLAANYSCAOFPRSAGYTELVKNGVEYGETREY-----GNFSFTRIYEGHEVP	495	
Db	404	GLGTQAAIDKLQLOBTSSWRWEEFD-SALGTUVVGVIYRKFKSGKGLFITVRGAGHMVP	462	
QY	496	YYQPIASIQLFNRTIFG	512	
Db	463	LWKPDSAFYMFKNFIDG	479	

Search completed: November 21, 2003, 18:00:07
Job time : 18 secs

This Page Blank (uspto)

Result No.	Score	Query %			DB	ID	Description
		Match	Length	No.			
1	894	30.0	423	2	S38953	carboxypeptidase D	
2	749	25.1	460	1	S51516	serine-type carbox	
3	652.5	21.9	1002	2	T43236	carboxypeptidase C	
4	636	21.3	508	1	S46008	probable carboxype	
5	601	20.2	552	2	JG7665	serine-type carbox	
6	596	20.0	542	1	JC1380	carboxypeptidase C	
7	594.5	20.0	523	1	S61713	carboxypeptidase C	
8	527.5	17.7	532	1	CPBYV	carboxypeptidase C	
9	473	15.9	429	1	T03607	probable carboxype	
10	455	15.3	500	1	S22530	carboxypeptidase C	
11	439	14.7	510	2	T48977	carboxypeptidase-1	
12	425.5	14.3	411	1	A35275	carboxypeptidase C	
13	417.5	14.0	2105	2	T18968	Probable serine-ty	
14	416.5	14.0	499	1	A29412	carboxypeptidase C	
15	383	12.9	482	2	T49079	serine-type carbox	
16	364	12.2	480	2	T50511	serine-type carbox	
17	359	12.1	487	2	T49080	serine-type carbox	
18	355.5	11.9	470	2	T16606	probable serine ca	
19	353	11.8	470	2	B96637	hypothetical prote	
20	351	11.8	482	2	A43828	Probable serine ca	
21	349	11.7	474	2	A35732	protective protein	
22	348	11.7	452	2	H84772	probable serine ca	
23	347.5	11.7	515	2	S44191	carboxypeptidase D	
24	344	11.5	465	2	G86244	Serine carboxypept	
25	344	11.5	501	2	T49081	serine-type carbox	
26	344	11.5	512	2	T33463	probable serine ca	
27	337.5	11.3	425	2	F85360	SERINE CARBOXYPEPT	
28	336	11.3	465	2	B85358	SERINE CARBOXYPEPT	
29	336	11.3	480	2	A31589	carboxypeptidase C	

Db 383 TVDNFSLKVVYAGHEVPPYQPDALQAFKQII-----QKK 418

RESULT 2

S51516

serine-type carboxypeptidase (EC 3.4.16.-) Z precursor - Absidia zychnae

C:Species: Absidia zychnae

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Jun-2000

C:Accession: S51516; S78013; S78014

R:Lee, B.R.; Takeuchi, M.; Kobayashi, Y.

Curr. Genet. 27, 159-165, 1995

A:Title: Molecular cloning and sequence analysis of the scp2 gene encoding the serine carboxypeptidase

A:Reference number: S51516; MUID:95308538; PMID:7789719

A:Accession: S51516

A:Molecule type: DNA

A:Residues: 1-460 <LEB>

A:Cross-references: EMBL:D16519; NID:5556466; PIDN:BAA03966.1; PID:9995456

A:Experimental source: strain NRIC 1199

A:Accession: S78013

A:Molecule type: protein

A:Residues: 52-62;90-99;367-381 <LEB>

A:Accession: S78014

A:Molecule type: mRNA

A:Residues: 18-460 <LES>

C:Genetics:

A:Gene: scp2

A:Introns: 64/3; 95/1; 107/2; 138/3; 234/2; 246/1; 270/1; 343/3; 353/2; 384/3; 437/3

C:Superfamily: serine carboxypeptidase

C:Keywords: glycoprotein; hydrolase; serine carboxypeptidase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-51/Domain: propeptide #status predicted <PRO>

F:52-460/Product: serine-type carboxypeptidase #status experimental <MAT>

F:40,46,119/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:186/Active site: Ser #status predicted

Query Match 25.1%; Score 749; DB 1; Length 460;

Best Local Similarity 35.7%; Pred. No. 4.4e-44;

Matches 179; Conservative 72; Mismatches 188; Indels 62; Gaps 15;

QY 12 LVAASWALPGSTPASVGRQLKPNPVGKLTITANNVIRYKPEAGVCEITPGVKYS 71

DB 18 LVCAPIVTVQAHMPSHLRQ-DGNDP-----SSGNTQDKYTP-----KLCD--PDVKQYS 66

QY 72 GYVDTSPESHFPFPEARHNPTAPITLWNGPGSDSLIGLFEELGPGCHVNSTPDDYI 131

DB 67 GLDAAANDHRYFPWFPEKNDPKNDPLTLWNGPGCSLLGLWEELGPGCQNGS---A 122

QY 132 NPHSMNEVSNLLFLSPLGPGVSYSDTVGSINPVTGVVENSFAGVQGRYPTIDATLID 191

DB 123 NPHSMHSSNMLFPDQPDGVGFSY-----GKQ-----TVS 152

QY 192 TTNLAEEAAWEILOGFLSPLDSRVQSKDFSLWTSYGGHYGPAFFNHFYQNBRIAN 251

DB 153 TTEDAARAWTFQAFVETFP-----QVSKLDVHYFGESYGGHYIPGFASHVDMNKKVQS 208

QY 252 GSVNGVQLNENSLGIINGIIDEAIOAPYYPEFAVNNYTGKAVNETVYNNMVFANOMPNG 311

DB 209 GEEKGVVPLKSLGVNGGFDAVIQYKSPKMTCHSTYPAVLGEE---ECDKMQQIYEND 265

QY 312 QDLISTCKGTNRTALADYALCAEATNMCRDNVEGYPYAGRGVYDIRHPYDDPTPSY 371

DB 266 CKPAEEQCAESD-----EDSDCVNANQC-GQIEG-IYAQSGYSFYDIRQQGDD-THPFF 317

QY 372 YNKFLAKDSYMDAIGVNNYNTQSNNDVYAFQOTGDFWPNFIEDLEELALPVRSLLIY 431

DB 318 VDE-LNKASVIVKEVGARGHSMCSDSVGTAFQTGCA-RSYIPAVEKLLKEGIPVLIYV 375

QY 432 GDADYICNWFPGQAVSLAANYSQAAQFRSAGYTPKLVNGVEYGHTRYGNFSTRVEAG 491

DB 376 GDADVICNWNGLDVADSLKWDGDAFSKTKLEAWKADKEVGQFRSADKLTFFVRVYEAG 435

QY 492 HEVPPYQPIASLQLPNRTIFG 512

Db 436 HEVPMYQPEAALSMFQTIWISS 456

RESULT 3

T43236

carboxypeptidase C (EC 3.4.16.5) precursor [validated] - fission yeast (Schizosaccharomyces)

N:Alternate names: carboxypeptidase Y

C:Species: Schizosaccharomyces pombe

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 08-Sep-2000

C:Accession: T43236; T37997

R:Tabuchi, M.; Iwihara, O.; Ohtani, Y.; Ohuchi, N.; Sakurai, J.; Morita, T.; Iwihara, J.

J. Bacteriol. 179, 4179-4189, 1997

A:Title: Vacuolar protein sorting in fission yeast: cloning, biosynthesis, transport, a

A:Reference number: Z22357; MUID:97352672; PMID:9209031

A:Accession: T43236

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1002 <TAB>

A:Cross-references: EMBL:D86560; NID:G3046860; PIDN:BAA25568.1; PID:G3046861

R:Olliver, K.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1997

A:Reference number: Z21760

A:Accession: T37997

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1002 <OLI>

A:Cross-references: EMBL:Z97209; PIDN:CAB10121.1; GSPDB:GNO0066; SPDB:SPAC19312.10C

A:Experimental source: strain 972h(-); cosmid c19G12

C:Comment: The mature carboxypeptidase exists as a heterodimer held together by a disul

C:Genetics:

A:Gene: cpy1; SPAC19312.10C

A:Map position: 1

C:Superfamily: serine carboxypeptidase

C:Keywords: disulfide bond; glycoprotein; heterodimer; hydrolase; serine carboxypeptidase

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-521/Domain: propeptide #status predicted <PRO>

F:200-332/Region: 13-residue repeats (H-H-[EK]-P-G-E-H-N-P-P-P-[MF])

F:344-425/Region: 9-residue repeats (E-H-H-[RKQ]-G-G-[KE]-[DE]-K)

F:522-1002/Product: carboxypeptidase C #status experimental <MAT>

F:627-880/Disulfide bonds: #status predicted

F:659/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:715,921,978/Active site: Ser, Asp, His #status predicted

Query Match 21.9%; Score 652.5; DB 2; Length 1002;

Best Local Similarity 33.8%; Pred. No. 5.9e-37;

Matches 165; Conservative 66; Mismatches 174; Indels 83; Gaps 15;

QY 49 TIRYKPEAGVCEITPGVKYSYGVYDTSFESHFPFPEARHNPTAPITLWNGPGS 108

DB 570 TIRVDSKPESLGIDT--VKQYTGILVDVEDRHLFEWFFESRNDPNDPVLWNGPGC 627

QY 109 DSIIGLFEELGPGCHVN-STFDDYINPHSMNEVSNLLFLSPLGPGVSYSDTVGSINPVT 167

DB 628 SSLTGLFMBELGPFSSINIELTKPEYNPHSMNSNASVFLDQPIINTGFSNGD--DSVLDTVT 695

QY 168 GVVENSSFAGVQGRYPTIDATLIDTNLAEEAAWEILOGFLSPLDSRVQSKDFSLWT 227

DB 686 -----AGKQVYAFNLNLEFAKFP---QYAHLDFFHAG 713

QY 228 ESYGGHYGPAFFNHFYQNER-----IANG-SVNGVOLNENSLGIINGIIDEAIOAPYYPE 282

DB 714 ESYAGHYIQQFAKEIMEHNGANFFVAGSYGMEKQVNLKSVLIGNLTDPLVQYFYGK 773

QY 283 FAVNNTYG-----IKAVNETVYNNMVFANOMPNGQDLISTCKGTNRTALADYAL 332

DB 774 MACESPYPIMSQEEDCRITGAYDT-----CAKLTGCVQTGFT-----PV 814

QY 333 CASATNMCRDNVEGYPYAFAGRGVYDIRHPYDDPTPSY-----YNNFLAKDSYMDAIGV 387

DB 815 CIGASLYCNAMIGP-FTKTLGLMIYDIRCECRQEHLCYPETGAIESYLAQEFVQEAIGV 873

QY 388 NINYQSNNDVYAFQOTGDFWPNFIEDLEEL--ALPVRVSLIY-GDADYICNWFPGQ 444

Db 874 EYDKGNTWVIGFLPKGDMWRKTRDDVTAILEAGLPV---LIYAGDADYICNYMGNE 930
QY 445 AVSLAANYSOAQRAGYTFPLKVGVEYGETREYGNFSTRVYVYAGHEVPYOPTASLQ 504
Db 931 AWTDALEWAGREFEYAEELKWPSPNGKEAGSGKSFNFGYLRLYEAGHMVFPNQPEASLE 990
QY 505 LFNRTIFG 512
Db 991 MLNSWIDG 998
RESULT 4
S46008
probable carboxypeptidase C (EC 3.4.16.5) YBR139w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YBR1015; hypothetical protein YBR139w
C:Species: Saccharomyces cerevisiae
C:Date: 26-Aug-1994 #sequence_revision 05-Jan-1996 #text_change 19-Apr-2002
C:Accession: S46008; S46581
R:Becam, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulecki, M.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45995
A:Accession: S46008
A:Molecule type: DNA
A:Residues: 1-508 <BEC>
A:Cross-references: EMBL:Z36008; NID:G536435; PIDN:CAA85097.1; PID:G536436; GSPDB:GN0000
A:Experimental source: strain S288C
R:Becam, A.M.; Cullin, C.; Grzybowska, E.; Lacroite, F.; Nasr, F.; Ozier-Kalogeropoulos,
Yeast 10(Suppl.A), S1-S11, 1994
A:Title: The sequence of 29.7kb from the right arm of chromosome II reveals 13 complete
A:Reference number: S46569; MUID:94378717; PMID:8091856
A:Accession: S46581
A:Molecule type: DNA
A:Residues: 1-508 <BE2>
A:Cross-references: EMBL:X75891; NID:G496856; PIDN:CAA53497.1; PID:G496869
A:Experimental source: strain S288C
C:Genetics:
A:Cross-references: SGD:S0000343
A:Map position: 2R
A:Note: MIPS:YBR139w
C:Superfamily: serine carboxypeptidase
C:Keywords: hydrolase; serine carboxypeptidase; transmembrane protein
F:6-22/Domain: transmembrane #status Predicted <TM>
F:219,415,474/Active site: Ser, Asp, His #status predicted
Query Match 21.3%; Score 636; DB 1; Length 508;
Best Local Similarity 32.6%; Pred. No. 3.2e-36;
Matches 167; Conservative 70; Mismatches 197; Indels 78; Gaps 14;
QY 19 LPGSTPASVGRQL-----PKNPTGVKTLTTANNVTIRYKEPGAEGVCETTFGVKSYGYV 74
Db 43 LPQNTQTLKLRNLNHDPLFTFTISSVDYTSRLRTVDPKSLGI----DTVKQWGYM 98
QY 75 DTPSESTFFFFFEARHNPTAPITLWNGPGSDSLGLFEELGVPCHVNSTFDDYINPH 134
Db 99 DYKDKHFFYFFFSRNDPANDPILLWNGPGGSSFTGLLFFELGPPSSIGADMKPIHPY 158
QY 135 SWNEVSNLLFLSQPLGVGFYSYDVTGDSINPVTGVVNSFAGVQGRYPTIDATLIDTIN 194
Db 159 SWNNASMIFFQLPVGFGSYGD-----EKVSTK 188
QY 195 LAEAANEILQGLFSLGSLDSRVQSDPSLWTSYGGHYGPAPFNHFYEQN-ERIANGS 253
Db 189 LAGDAVIFELFFTEAFPHL-----RSNDFHAGESYAGHYIPQIAHEIVKNER----- 239
QY 254 VNGVQLNFNSLGIINGIIDEAIOAPYPEFAY-NNTYGIKAVNETVNYMKFANQMGNC 312
Db 240 ----TFNLTSVWINGITDPLIQADYEPMACGGHGHVPLVSSBECEKMSKAAGR----- 290
QY 313 QDLISTCKQTNRTALADYAL--CAEATNMCNDVNGVYAFAGVGVDIRHPYDDP---- 366
Db 291 -----CRRLNKLKCYASKSSLPICVATAYCDSALLEIYIN-TGLNVDIRGPCEDNSTDG 343
QY 367 ---TPPSYNNKFLAKDSVMDAIGVNI-NYTSNNNDVYAFQQTGDFVWPNFIEDLEEILA 422

Db 344 MCVTGRYVDQYNNMFVEQETLGLSDVHNSGCDNVFTGLFTDGSKP-FQQYIAELLN 402
QY 423 LPVRVSLIYGDADYICNWFQGGQAVSLAANYSOAQRFSAGYTP--LKVNGVEYGETREYG 480
Db 403 HNIPLVLIYAGDKDYICNWLGNHANSLEWINKRRIQRMRLRPWVSKETGEELGQVKNYG 462
QY 481 NFSFTRYVYAGHEVPYQYPIASLQLFNRTIFG 512
Db 463 PFTFLRIYDAGHMVYPYDQPEASLEMVNSWISG 494
RESULT 5
JC7666
serine-type carboxypeptidase homolog precursor - Emeritella nidulans
N:Alternate names: carboxypeptidase Y homolog
C:Species: Emeritella nidulans, Aspergillus nidulans
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C:Accession: JC7666
R:Ohsumi, K.; Matsuda, Y.; Nakajima, H.; Kitamoto, K.
BioSci. Biotechnol. Biochem. 65, 1175-1180, 2001
A:Title: Cloning and characterization of the cpYA gene encoding intracellular carboxypep
A:Reference number: JC7666; MUID:21333188; PMID:11440134
A:Accession: JC7666
A:Molecule type: mRNA
A:Residues: 1-552 <OHS>
A:Cross-references: DBJ:AB051820
C:Comment: This protein is an intracellular carboxypeptidase, which is localized in vacu
F:131-132/Region: cleavage site, by Arg-Ile #status predicted
C:Genetics:
A:Gene: cpYA
A:Introns: 165/1
F:1-15/Domain: signal sequence (or 1-18, or 1-19) #status predicted <SIG>
F:16-131/Domain: pro-sequence, vacuolar targeting signal (or 19-131, or 20-131) #status
F:131-132/Region: cleavage site, by Arg-Ile #status predicted
Query Match 20.2%; Score 601; DB 2; Length 552;
Best Local Similarity 31.6%; Pred. No. 9.3e-34;
Matches 154; Conservative 77; Mismatches 175; Indels 82; Gaps 19;
QY 45 ANNVTIRYKEPGAEGVCETTFGVKSYGVVDTSP-ESHTFFFEARHNPTAPITLWLN 103
Db 127 AVDLRIKKTDFSSLGI---DDVKQYTGYLQDNENKELFWFFESRNDKNDPVPVLMN 183
QY 104 GPGSDSLIQLFEELGVPCHVNSTFDDYINPHSWNEVSNLLFLSQPLGVGFYSYDVTGDSI 163
Db 184 GPGGSSLTGLFELGPPSSIDENIKPVYNPVAVNSASVIFLDQPVNVGYSYS----- 236
QY 164 NPVTGVVNSFAGVQGRYPTIDATLIDTTLAAEAANEILQGLFSLGSLDSRVQSKDF 223
Db 237 -----GSTVSDTV-AAGKDVYVALLTLFFKQFP-----EYAEQDF 269
QY 224 SLWTSYGGHYGPAPFNHFYEQNERIANGSVNGVQLNFNSLGIINGIIDEAIOAPYPEF 283
Db 270 HIAGESYAGHYIPVFTSEILSHQR-----NINLKSVLINGITDGTQVYEYRPM 320
QY 284 AVNNTYGIKAVNETVNYMKFANQMPNG---CQDLISTCKQTNRTALADYALCAEATNMC 340
Db 321 ACSEGG-GYPAVLDE-----SSCRSMDNALGRCSQMIESCYNSESA-----WVCVPASTYC 369
QY 341 RDNVGEGYPAAGVGVDIRHPYDDPT-----PPSYNNKFLAKDSVMDAIGVNI-N-YTQS 394
Db 370 NNALLAPYQR-TGQNVYDVGKCEDESNLCYKMGVYVSEYLNKPEVRAAAGVGDVDS 428
QY 395 NNDVYAFQQTGDFVWP--NFIEDLEEILALFVRVSLIY-GDADYICNWFQGGQAVSLAAN 451
Db 429 NFDLNRNFLFHGDWMEYHRLVPGILE--QIPV---LIYAGDADFCINLGNKAWTSALE 483
QY 452 YSOAQRFSAGYTPLKV-----NGVEYGETREYGNFSTRVYVYAGHEVPYQYPIASLQ 506
Db 484 WPGHKEFAAAPMEDLKLVDNEHTGKIGQIKTHGNFTFMRLYGGHVMVMDQPEASLEFF 543
QY 507 NRTIFG-W 513

carboxypeptidase C (EC 3.4.16.5) precursor - yeast (Pichia pastoris)
N;Alternate names: carboxypeptidase Y
C;Species: Pichia pastoris
C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 16-Jun-2000
C;Accession: S61713
R;Ohl, H.; Ohtani, W.; Okazaki, N.; Furuhashi, N.; Ohmura, T.
Yeast 12, 31-40, 1996
A;Title: Cloning and characterization of the Pichia pastoris PRC1 gene encoding carboxy-
peptidase Y
A;Reference number: S61713; MUID:96381245; PMID:8789258
A;Accession: S61713
A;Molecule type: DNA
A;Residues: 1-523 <OHI>
A;Cross-references: EMBL:X87987; NID:g1171615; PIDN:CAA61240.1; PID:g1171616
C;Superfamily: serine carboxypeptidase
C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase; zymogen
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-106/Domain: propeptide #status predicted <PRO>
F;107-522/Product: carboxypeptidase C #status predicted <MAT>
F;193,271,484,487/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;249,441,498/Active site: Ser, Asp, His #status predicted
F;296-310,327-336/Disulfide bonds: #status predicted

Query Match 20.0%; Score 594.5; DB 1; Length 523;
Best Local Similarity 33.1%; Pred. No. 2.4e-33;
Matches 153; Conservative 59; Mismatches 169; Indels 81; Gaps 16;

QY 68 KSYSGYDT-SPESTHTFWPEARHNPETAPITLWNGPGSDSLGLFEELGPHVNST 126
DB 121 KQYSGYLDVEADKHFYWFESRNDPQNDPIILWNGPGSSLTGLFELGSSRINEN 180
QY 127 FDDYINSHWNEVSNLLFLSOPLGVSFSTVDGSIINFTGVVENSFAGVQGRYPTID 186
DB 181 LKPIFNPSWNGNIIYLDQPVNVGFSYS----- 210
QY 187 ATLIDTTLAAEAAWEILQGLSLPSLDSRVQSKDFSLMTESYGHGYPAPFNFYQON 246
DB 211 SSSVSNTVAGEDVYAFLOLFFQHP-----EYQTNDFHAGESYAGHYIPVFAEDILSQK 266
QY 247 ERIANGSVNGVLNENSLGII--NGIIDEAIQAPYYPFAVNTYGIKAVNETVYNYMKF 304
DB 267 NR-----NENLTSVLINGLTDPLQYRYEPMACGEG-GAPSVLP----- 306
QY 305 ANQMN-----GCQDLISTCKQTNRTALADYALCAEATNMCRDNVEGYAFAGRGVYD 358
DB 307 ADECECNMLVTQDKLSLIQACVDSQA-----FTCAPAIYCNAQMGPEYQR-TGKNVYD 360
QY 359 IRHPYDDPT-----PPSYNKFSLAKDSVMDAIGVNI--YTQSNNDVYAPQQTGDFWENF 413
DB 361 IRKEDGGSLCYKDLFTDYLNQKFQDQALGAEDVTYESCNFEINRNLFFAGDWMKP-Y 419
QY 414 IEDLEIL--ALPVRVSLIY-GDADYICNWFQGVQAVSLAANYSQAAQFERSAGYTPLKVNG 470
DB 420 HEHVSLLNKGLEP---LLIYAGDKDFICNWLGNRAWTDLVFWDDAGFEKAEVQDWLVNG 476
QY 471 VYGYETREYGNFSFTRVYEAAGHEVPYQPIASLQLENRTIFG 512
DB 477 RKAGEFKYNTLYLVYDAGEWAPYDQENSHENVRWISG 518

RESULT 8
CPBY
carboxypeptidase C (EC 3.4.16.5) precursor [validated] - yeast (Saccharomyces cerevisiae)
N;Alternate names: carboxypeptidase Y; protein YMR297W
C;Species: Saccharomyces cerevisiae
C;Date: 15-Oct-1982 #sequence_revision 31-Mar-1993 #text_change 15-Sep-2000
C;Accession: A26597; S47458; A94609; A00909
R;Valls, L.A.; Hunter, C.P.; Rothman, J.H.; Stevens, T.H.
Cell 48, 887-897, 1987
A;Title: Protein sorting in yeast: the localization determinant of yeast vacuolar carb-
oxypeptidase Y
A;Reference number: A26597; MUID:87131100; PMID:3028649
A;Accession: A26597
A;Molecule type: DNA
A;Residues: 1-532 <VAL>

carboxypeptidase C (EC 3.4.16.5) precursor - yeast (Candida albicans)
N;Alternate names: carboxypeptidase Y
C;Species: Candida albicans
C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999
C;Accession: JCI1380
R;Mukhtar, M.; Logan, D.A.; Kaeufer, N.F.
Gene 121, 173-177, 1992
A;Title: The carboxypeptidase Y-encoding gene from Candida albicans and its transcriptio
n
A;Reference number: JCI1380; MUID:93051356; PMID:1427093
A;Accession: JCI1380
A;Molecule type: DNA
A;Residues: 1-542 <MUK>
A;Cross-references: GB:M95182; NID:g170828; PIDN:AAA34326.1; PID:g170829
A;Note: the authors translated the codon GAT for residue 42 as Asn, AAA for residue 373
C;Genetics:
A;Gene: CPY1
C;Superfamily: serine carboxypeptidase
C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-127/Domain: propeptide #status predicted <PRO>
F;128-542/Product: carboxypeptidase C #status predicted <MAT>
F;213,291/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;269,461,518/Active site: Ser, Asp, His #status predicted
F;316-330,347-356/Disulfide bonds: #status predicted

Query Match 20.0%; Score 596; DB 1; Length 542;
Best Local Similarity 31.5%; Pred. No. 2e-33;
Matches 169; Conservative 64; Mismatches 193; Indels 110; Gaps 20;

QY 19 LPGSTP--ASVGRRLPKNPTGVKLTANNVT--IRYKEPAGVCEY----- 63
DB 71 LDGLTPEIKNIWLMKFP-----NSITLNFRAKPKKGIITQQDFHVTDAQV 121
QY 64 -----TP-----GVKSYGYVDTSP--SHTFWFFFEARHNPETAPITLWNGPGF 107
DB 122 PNHKLRISTPKDLGIDTVKQSGVLDVVDDEKHFYFFESRNDPKNDPVLWNGPGF 181
QY 108 SDSLGLPEELGPHVNSTFDDYINPHSWNEVSNLLFTSOPLGVSFSDTVGDSINPVT 167
DB 182 CSSLGLPEELGPHVNSTFDDYINPHSWNEVSNLLFTSOPLGVSFSDTVGDSINPVT 230
QY 168 GVENSFPAGVQGRYPTIDATLIDTTLAAEAAWEILQGLSLPSLDSRVQSKDFSLWT 227
DB 231 -----SQSVSNTIARAGKDYAFLOLFFKNP-----EYANLDFHIAG 267
QY 228 ESYGHHYGAFFNHFYEQNERIANGSVNGVQLNENSLGII--NGIIDEAIQAPYYPFAV 285
DB 268 ESYAGHYIAFASEILITHEP-----NENLTSVLINGLTDPLVQVEYYPMAC 316
QY 286 NNTYGIKAVNETVYNYMKFANQMPNGCQDLISTCKQTNRTALADYALCAEATNMCRDNVE 345
DB 317 GEGGEPVLEPEECQGM--LNLPLR-CLSLIESCYE-----GSVMSCVPATIIYCNQGM 368
QY 346 GPYAFAGRGVYDIRPHYDDP-----TPPSYNKFSLAKDSVMDAIGVNI--YTQSNNDVY 400
DB 369 GPYQK--TGRNVYDIRTWCEGSSLCYQLEVIDYIYLNLPVKALGAEVDEYQSCNFDIR 427
QY 401 AFQQTGDFVMP---NFIEDLEILALPVRVSLIY-GDADYICNWFQGVQAVSLAANYSQAA 456
DB 428 NFMFAGDWMKPYQKNVLDLLEK--ELPV---LLIYAGDKDFICNWLGNQAWTNLEWSGSK 482
QY 457 QFRSAGYTPLVKNGVEGETREYGNFSFTRVYEAAGHEVPYQPIASLQLENRTIFG 512
DB 483 GFTKAPKSWKVGKNAAGEVKNYKHFTFLRVFGGHHMVPYDQENALDMWNRWISG 538

RESULT 7
S61713

A;Cross-references: EMBL:M15482; NID:g172238; PIDN:AAA34902.1; PID:g172239
submitted to the EMBL Data Library, August 1994
A;Reference number: S47445
A;Accession: S47458
A;Molecule type: DNA
A;Residues: 1-532 <BAR>
A;Cross-references: EMBL:X80836; NID:g1289327; PIDN:CAA56806.1; PID:g530354; GSPDB:GN000
R;Svendsen, I.; Martin, B.M.; Viswanatha, T.; Johansen, J.T.
Carlsberg Res. Commun. 47, 15-27, 1982
A;Title: Amino acid sequence of carboxypeptidase Y. II. Peptides from enzymatic cleavage
A;Reference number: A90763
A;Accession: A90763
A;Molecule type: protein
A;Residues: 112-223, 'X', 225; 228-239, 'X', 241-259, 'HG', 262-267, 'X', 269-388, 'E', 390-451, 'N'
R;Svendsen, I.
submitted to the Atlas, October 1982
A;Reference number: A94609
A;Contents: disulfide bond
A;Accession: A94609
A;Molecule type: protein
A;Residues: 224-227 <SV2>
A;Note: this is a revision to the sequence in reference A90763
R;Martin, B.M.; Svendsen, I.; Viswanatha, T.; Johansen, J.T.
Carlsberg Res. Commun. 47, 1-13, 1982
A;Title: Amino acid sequence of carboxypeptidase Y. Peptides from cleavage with cyanogen
A;Reference number: A90762
A;Contents: annotation; experimental details
R;Endrizzi, J.A.; Remington, S.J.
submitted to the Brookhaven Protein Data Bank, March 1994
A;Reference number: A52472; PDB:1YSC
A;Contents: annotation; X-ray crystallography at 2.8 angstroms; disulfide bonds
C;Genetics:
A;Gene: SGD:PRC1; LBC1; MIPS:YMR297w
A;Cross-references: SGD:S0004912; MIPS:YMR297w
A;Map position: 13R
C;Superfamily: serine carboxypeptidase
C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase; yeast vacuole; zymogen
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-111/Domain: propeptide #status predicted <PRO>
F;112-532/Product: carboxypeptidase C #status experimental <MAT>
F;124,198,279/Binding site: carboxydrate (Asn) (covalent) #status predicted
F;167-409,328-351,335-344,373-379/Disulfide bonds: #status experimental
F;257/Active site: Ser #status experimental
F;449,508/Active site: Asp, His #status predicted
Query Match 17.7%; Score 527.5; DB 1; Length 532;
Best Local Similarity 29.1%; Pred. No. 1.1e-28;
Matches 149; Conservative 71; Mismatches 195; Indels 97; Gaps 19;
QY 33 PKNPTGVKT-----LTTANNVIRY-----KEPGAEGVCBTRPGVKSYGVVDTSP- 79
DB 82 PRFEPAIKTKDMDFFVKNDALENQLRVNKKDKPKLGI---DENVTOYTGVLDEVED 138
QY 80 SHTFWFPEARINPETAPITLWNGPGSDSLIGLFEELGPGCHVNSTFDDYNPHSWNEV 139
DB 139 KHFFWFTEESRNDPAKDPVILWNGPGGCSLTGLFFELGPGSSIGPDLKIPGNPYSWNS 198
QY 140 SNLLFLSOPLGVSFSDVDSINPVTGVVNSFAGVQGYPTIDATLITDITNLAABA 199
DB 199 ATVIFLDQPVNVGFSYS-----GSSG-----VSNVAAAGKD 229
QY 200 AWEILQGLSLGSLDSRVQSKDFSLWTSYGGHYGPAFFNHFYQNERIANGSVNGVL 259
DB 230 VYNFLELFDQPEYVKNQ--DFHIAGESYAGHYIPFASILSHKOR----- 276
QY 260 NPNLSGII--NGIIDEAIQAPYYPEFA-----VNNYGIKAVNETVYNTMKFANQP 309
DB 277 NFNLTSVLIGNLTDLPTQYNYEPWACGEGPSVLPSECSAMEDSL----- 325
QY 310 NGCQDILSTCKQNTALADYALCAEATNMCRDNVEGYPYAFAGRVYDIRHPYDD----- 365

DB 326 ERCLGLIESCYDSQ-----SWMSCVPATIIYCNNAQLAPYOR-TGRNVYDIRKCEGNNLC 379
QY 366 -PTPPSYNFKLAKDSVMDAIGVNNINTQS--NNDVYVAFQQTGDFVWPNFIEDLEEIIAL 423
DB 380 YPTLQD-IDDYLNQDYVKEAVGAEDVHYESCNFDINENFLFAGDWMKP-YHTAVTDLLNQ 437
QY 424 PVRVSLIYGDADYICNVWFGQAVSLAANYSQAAQFRS---AGYTPLKVNVEGVEGETREYG 480
DB 438 DLPILVYAGDKDFICNWLGNKAWTDVLPWKYDEFEASQKVRNWTASITDEVA-GEVRSYK 496
QY 481 NFSFTRVYAGHEVPYVQPIASLQLENRTIFG 512
DB 497 HFTYLVFNGGHMVPDPENALSMVNEWIHG 528
RESULT 9
T03607
probable carboxypeptidase C (EC 3.4.16.5) cbp31 - rice
N;Alternate names: serine-type carboxypeptidase homolog
C;Species: Oryza sativa (rice)
C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 21-Jul-2000
C;Accession: T03607
R;Washio, K.; Ishikawa, K.
Plant Physiol. 105, 1275-1280, 1994
A;Title: Organ-specific and hormone-dependent expression of genes for serine carboxypept
A;Reference number: Z14975; MUID:95062718; PMID:7972496
A;Accession: T03607
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-429 <WAS>
A;Cross-references: EMBL:D17587; NID:g409581; PIDN:BAA04511.1; PID:g409582
A;Experimental source: cv. Yukihihikari
C;Genetics:
A;Gene: cbp31
C;Superfamily: serine carboxypeptidase
C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase
F;76,414,417/Binding site: carboxydrate (Asn) (covalent) #status predicted
F;148,336,393/Active site: Ser, Asp, His #status predicted
Query Match 15.9%; Score 473; DB 1; Length 429;
Best Local Similarity 28.6%; Pred. No. 4.5e-25;
Matches 130; Conservative 77; Mismatches 178; Indels 70; Gaps 16;
QY 70 YSGY--VDTSPSRHTFFWFPEARHNP-ETAPITLWNGPGSDSLIGLFEELGPGCHVNST 126
DB 17 HAGYVRLPNTHDARLFYFFESKSGKGGDDPVVILWLTGGPGCSSELALFYENGPRPHIADN 76
QY 127 FDDYNPHSWNEVSNLLFLSOPLGVSFSDVDSINPVTGVVNSFAGVQGYPTID 186
DB 77 MSLVNDPFGWDESNLIIVDQGTGTFYSNPRDTRHDEAG-VSNDLYA----- 125
QY 187 AFLIDITNLAABAWEILQGLSLGSLDSRVQSKDFSLWTSYGGHYGPAFFNHFYQNER 246
DB 126 -----FLQAPFTEHPNF-----AKNDFYITGESYAGHYIPASRVYKGN 165
QY 247 ERIANGSVNGVQLNFSNGIINGIIDEAIQAPYYPEFAVNNYTGKAVNETVYNTMKFAN 306
DB 166 K-----NSEGHINILKGFAGNGLTDPATQVKAITYDSLD-----MGLTKSQFNRI--N 213
QY 307 QMPNGCQDILSTCKQNTNR-TALADYALCAEATNMCRDNVEGYPYAFAGRVYDIRHP 362
DB 214 KIVPTCELAIKLCTSGTISCLGAYVVC-----NLICSSIE---TIIGKKNYDIRKPCVG 266
QY 363 ---YDDPTPPSYNFKLAKDSVMDAIGV-NINYTQSNNDVYVAFQQTGDFVWPNFIE-DL 417
DB 267 SLCYD-----LSNMEKFLQKSVRESLGVGDITQFVSCFTVIQAML-----LDWMRLEVGI 318
QY 418 EEILALPVRVSLIYGDADYICNVWFGQAVSLAANYSQAAQFRSAGYTPLKVNVEGVEGETR 477
DB 319 PELLENDIKVLIYAGEYDLICNWLNSRWVNSMWSGKEAFVSSSEEFPTVDGKEAGILK 378
QY 478 EYGNFSFTRVYAGHEVPYVQPIASLQLENRTIFG 512

Db 379 SYGPLSFLKVDHAGHVMVPMQPKVALEMLMRWTSG 413

RESULT 10

S22530 carboxypeptidase C (EC 3.4.16.5) precursor - rice

N;Alternate names: carboxypeptidase III

C;Species: Oryza sativa (rice)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 24-Nov-1999

C;Accession: S22530

F;Washio, K.; Ishikawa, K.

Plant Mol. Biol. 19, 631-640, 1992

A;Title: Structure and expression during the germination of rice seeds of the gene for a

A;Reference number: S22530; MUID:92329723; PMID:162776

A;Accession: S22530

A;Molecule type: DNA

A;Residues: 1-500 <WAS>

A;Cross-references: EMBL:S40458

C;Genetics:

A;Introns: 100/2; 156/3; 196/3; 244/3; 295/1; 324/3; 411/1; 455/3

A;Superfamily: serine carboxypeptidase

C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase; zymogen

F;1-21/Domain: signal sequence #status predicted <SIG>

F;122-73/Domain: propeptide #status predicted <PRO>

F;74-484/Product: carboxypeptidase C #status predicted <MAT>

F;485-499/Domain: carboxyl-terminal propeptide #status predicted <CPRO>

F;144/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;216,404,461/Active site: Ser, Asp, His #status predicted

Query Match 15.3%; Score 455; DB 1; Length 500;

Best Local Similarity 26.4%; Pred. No. 9.8e-24;

Matches 138; Conservative 85; Mismatches 225; Indels 74; Gaps 17;

QY 7 LSVLPVLAASMA---LP--GSTPASVGR-----QLPK--NPTGVKTLTANNVTIR 52

Db 10 LUVVVLASACAGELRPDAKPAQAERLRSNLLPKKAGPTGAGVPSVAGELLE 69

QY 53 KPGAGGVCETTFGVKSYSGY--VDTPSPESHFFWPFARHNPETAPITLWNGFGSDS 110

Db 70 RRVTLPLGQGVGDGLGHAGYVRLPNTHDARFYLFSRGGKED--PVVIMLTGGPGCSS 128

QY 111 LGLFEEELGPCVHNSTFDDYINPHSNWNSNLLFSLQPLGVGFSYSDTVGSGINPTGVV 170

Db 129 ELAVFYENGPEFTISNNMSLAWNKFGWDTISNIFVDQPTGTGFSYSSDRDRTRHDETGV- 187

QY 171 ENSSPFAGVQGRYPTIDATLIDITNLAAEAAWEILOGLSLPSLDSRVQSKDFSLWTSY 230

Db 188 -----SNDLYSFLQVFFKKHDEF-----AKNDFFITGESY 217

QY 231 GGHYGPAPFNHFYQNERIANGSVNGVQLNFSGLINGIIDEAIQAPYYPFPAVNNYTG 290

Db 218 AGHYIIPAFASRVHOGNK--AN---EGHINLKGFAINGLTDPAIQKAYTDYALD--- 268

QY 291 IKAVNETVYNYMKFANQMGCCODLSTCKQTNRTALADYALCAEATNCRDNVEGPPYA 350

Db 269 MNLIKKSDYDRI---NKFTPPCFALKCTGNK-----ASCMAYMVNCSIFSIMKL 319

QY 351 FAGRGVYDIRHPYDDPTPPSYN--KFLAKDSVMDAIGV--NINYTQSNNDVYVAFQOTGD 407

Db 320 VGTKNYDVRKCEGKCLYDFSNLEKFFGDKAVKEAIGVGDLEFVSCSTTVYQAMLTJ-- 377

QY 408 FVWPNFIE--DLEILALPVRVSLIYGDADYICNWFQGVQAVSLAANYSOAAQFRSAGYTPL 466

Db 378 --WNRNLEVGIPALLEDGINLVYAGEYDLICNWLGNRSRVHMEWSGQKDFVSSHPF 435

QY 467 KUNGVGEFTREYGNFSFTVYAGHEVPYQPIASIQLENRTIFG 508

Db 436 VVDGAEAGVLKSHGFLPSFLKVHNAHVMVPMQPKASLEMLRR 477

RESULT 11

T48977

carboxypeptidase-like protein Fl4D17.80 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 24-Oct-2000

C;Accession: T48977

R;Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Rudd, S.;

submitted to the Protein Sequence Database, April 2000

A;Reference number: 225008

A;Accession: T48977

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-510 <JOR>

A;Cross-references: EMBL:AL353992; GSPDB:GN00061; ATSP:FL4D17.80

A;Experimental source: cultivar Columbia; BAC clone F14D17

C;Genetics:

A;Gene: ATSP:FL4D17.80

A;Map position: 3

A;Introns: 115/2; 171/3; 211/3; 259/3; 310/1; 341/3; 428/1; 472/3

C;Superfamily: serine carboxypeptidase

F;231,421,478/Active site: Ser, Asp, His #status predicted

Query Match 14.7%; Score 439; DB 2; Length 510;

Best Local Similarity 26.6%; Pred. No. 1.3e-22;

Matches 123; Conservative 81; Mismatches 189; Indels 69; Gaps 14;

QY 64 TPGVKSY---SGY--VDTPSPESHFFWPFARHNPETAPITLWNGFGSDSLIGLFFEL 118

Db 93 SPVQDFGHHAGYVYKLPNSKARMFYFFESRTN--KADPVVIMLTGGFGCSSELALFYEN 151

QY 119 GPCVHNSTFDDYINPHSNWNSNLLFSLQPLGVGFSYSDTVGSGINPTGVVENSFPAGV 178

Db 152 GPFTVSNSSLSWNEFGWMDKASNLTYVDQPTGTGFSYTSQSDLRHDEGTV----- 202

QY 179 QGRYPTIDATLIDITNLAAEAAWEILOGLSLPSLDSRVQSKDFSLWTSYGGHYGPAF 238

Db 203 -----SNDLYDLOAFPEKHP-----QFVNDFYITGESIAGHYIPAL 240

QY 239 FNHFYQNERIANGSVNGVQLNFSGLINGIIDEAIQAPYYPFPAVNNYTGKAVNETV 298

Db 241 ASRVHGNK-----NKEGTHINLKGFAINGLTFEIQYGAADYALDNLITQSDHNL 295

QY 299 YNYMKFANQMGCCODLSTCKQTNRTALADYALCAEATNCRDNVEGPPYAPAGR- 354

Db 296 NRY-----YATCQSIKECSADGEGDCASSYTVG--NNIFQKIMDIAGNV 340

QY 355 GYVDIRHPYDDPTPPSYN--KFLAKDSVMDAIGV--NINYTQSNNDVYVAFQOTGDVFWP 411

Db 341 NTYDVRKQCEGSLCYDFSNMENFLNOKSVRKALGVGDIEFVSCSTAVIYFAMQND---- 396

QY 412 NFIE--DLEILALPVRVSLIYGDADYICNWFQGVQAVSLAANYSOAAQFRSAGYTPLKVG 470

Db 397 RNLEVGIPALLOGGKILLVYAGEYDLICNWLGNRSRVHMEWSGQKEFVAATVPFPHVDN 456

QY 471 VNYGTEYGNFSFTVYAGHEVPYQPIASIQLENRTIFG 512

Db 457 KEAGLMKNYGSITFLKVDHAGHVMVPMQPKAALQMLQNMVQ 498

RESULT 12

A35275

carboxypeptidase C (EC 3.4.16.5) - barley

N;Alternate names: carboxypeptidase III

C;Species: Hordeum vulgare (barley)

C;Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999

C;Accession: A35275

R;Sorensen, S.B.; Svendsen, I.; Breddam, K.

Carlsberg Res. Commun. 54, 193-202, 1989

A;Title: Primary structure of carboxypeptidase III from malted barley.

A;Reference number: A35275; MUID:90315015; PMID:2639682

A;Accession: A35275

A;Molecule type: protein

A;Residues: 1-411 <SOR>

C;Superfamily: serine carboxypeptidase

C;Keywords: blocked amino end; glycoprotein; hydrolase; serine carboxypeptidase

F11/Modified site: blocked amino end (Leu) (probably acetylated) #status experimental
 F71/Binding site: carboxylate (Asn) (covalent) #status experimental
 F143,331,388/Active site: Ser, Asp, His #status predicted

Query Match	14.3%;	Score 425.5;	DB 1;	Length 411;
Best Local Similarity	26.5%;	Pred. No. 8.1e-22;		
Matches 121;	Conservative 71;	Mismatches 203;	Indels 62;	Gaps 12
Qy	59	GVCETTPGVKYSGY--VDTPSESHTFWFFFEARHNPTAPITLLWNGPGSDSLIGLIFE	116	
Db	3	GLPEGVADLGHAGYRLNTHDARMFYFFESRGKED-PWVILWTGGPCSSSELAVFY	61	
Qy	117	ELGPCHVNSTDDYINPHSMWEVSNLLFLSQPLGVGFYSYSDTVGSGINPTVGTVWENSFA	176	
Db	62	ENGFTIANNMSLVNKGWDKISIIIFVDQPTGFGFYSDDRRDTRHDETV-----	114	
Qy	177	GVQGRYPTIDATLIDTTLNLAEEAAWEIIQGLFSLGFLSDRSVQSKDFSLWTESYGHYGP	236	
Db	115	-----SNDLYDFLOVFFKKHPEF-----IKNDFFITGESYAGHYIP	150	
Qy	237	AFNFHFYEQNERIANGSVNGVLNFSLGIINGIIDEAIOAPYYPEFAVNNTYGIKAVNE	296	
Db	151	AFASRVHQGNKK-----NEGTHINLKGFALNGLTDPALXYKATDYALENNLLQKDAYE	205	
Qy	297	TVVYMKFANPMNGCCDLISCTQTVR-TALAYALCAEATNMCNRNVEGYYAFAGRG	355	
Db	206	RI-----NKFTPPCCFAIKLCCTNGKASCMAAVYVNCNTIFNSIMKLVTGKNY-----	252	
Qy	356	VYDIRHPYDPTPPSYN--KFLAKDSVMDAIGV-NINYTCSDNDVYYAFQQTGDFWPN	412	
Db	253	-YDVRKECEGKLCYDFSNLSEKFFGDKAVRAQIAGVGDFIEFVSCSTSVYQAMLT-----WMR	307	
Qy	413	FIE-DLEETILALPVRVSLIYGDADYICNWFGGQAVSLAANYSOAAQFSAGYPLKTVNGV	471	
Db	308	NLEVGIPALLEDGINVLIIYAGEYDLICNWLNSRWHSMEWSGXQFAKTAESSFLVDDA	367	
Qy	472	EYGETREYGNPSTRVYVYAGHEVPYYQPIASLQLFNR	508	
Db	368	QAGVLYKSHGALSFLKVNHAGHMVYEMDOPKAALEMIR	404	

RESULT 13

T18968
probable serine-type carboxypeptidase (EC 3.4.16.-) Y16B4A.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Mar-2000
C/Accession: T18968; T23145; T26477
R/Thomas, K.
submitted to the EMBL Data Library, March 1996
A/Reference number: Z19053
A/Accession: T18968
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-2105 <WIL>
A/Cross-references: EMBL:Z70203; PIDN:CAA94110.1; GSPDB:GN000028; CESP:Y16B4A.2
A/Experimental source: clone C05G5
R/Ilyovd, C.
submitted to the EMBL Data Library, March 1997
A/Reference number: Z19697
A/Accession: T23145
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-2105 <W12>
A/Cross-references: EMBL:Z93339; PIDN:CAB07544.1; GSPDB:GN000028; CESP:Y16B4A.2
A/Experimental source: clone H40D05
R/Wallis, J.
submitted to the EMBL Data Library, June 1998
A/Reference number: Z20220
A/Accession: T26477
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-2105 <W13>
A/Cross-references: EMBL:AL023825; PIDN:CAA19443.1; GSPDB:GN000028; CESP:Y16B4A.2

A;Experimental source: clone Y16B4A
C;Comment: This protein has 4 repeats of sequence homologous to serine carboxypeptidase
C;Genetics:
A;Gene: CESP:Y16B4A.2
A;Map position: X
A;Introns: 43/3; 67/2; 179/1; 208/3; 239/2; 281/2; 296/3; 346/2; 444/3; 503/3; 588/3; 616/4/3; 1752/2; 1794/3; 1834/3; 1882/2; 1905/2; 1942/3; 1982/3
C;Keywords: duplication, hydrolyase, serine carboxypeptidase

Query Match	14.0%;	Score 417.5;	DB 2;	Length 2105;
Best Local Similarity	23.5%;	Pred. No. 2.8e-20;		
Matches 156;	Conservative 91;	Mismatches 209;	Indels 209;	Gaps 28;
Qy	2	RGYEFLSVLPVUASWA-----LPSTPASVGRRLPKNPFGVKTLTTANN 47		
Db	1467	QGSQYVTIIAGYAKSWTQNLVQLTVKSGSHFVPSDRPAQ-----ALQMLTNFLS 1515		
Qy	48	VTIRYKPGAGV-----CET-----TPGV-----KSYSGYVDTSPE 80		
Db	1516	NQANYSTPAGIDVTPQLTNWISQTNCTTQCTRIINLPGLPADMLFKQSGFLDGLSGH 1575		
Qy	81	HTPFWFPEARHNPETAPITLWLGPGSDSLIGLFEELGPHCHVNSTEDDYI-NPHSMNEV 1339		
Db	1576	KVHYWLVESENPNSTDPLLLWNLGPGSSLSMGLFEENGFFRKQSDQLSRNPYSWNKF 1635		
Qy	140	SNLLFLSQPLGVGFSDYTDVGSINPVTGVVENSFAGVQGRYPTIDATLIDTNNLAAEA 199		
Db	1636	ANVLYESPIGVGSYA-----YNNNTNIQYDDVT--TAQE 1668		
Qy	200	AWELLQGLSGLPSLDSRVQSKDFSLWTESYGHYGFAPFNHFEYQNERIANGSVNG-VQ 258		
Db	1669	NYAALKSFFAQYP---QYTTSDFYTTGESYAGVYLPGL-----SALLVQGIKSGDIN 1717		
Qy	259	LNFNLSGIINGIIDE-----AIQAPYYPEFVANNYGIKAVNNTVYN-----300		
Db	1718	INYKGVSIGNGVIDKRTDMNSQLHYQY-----HG-GISATTYNTALQLCCSGDEF 1767		
Qy	301	-----YMKFANQMENG-----CQD-LISTCKQTNRTALADYALCAEATNMCRDNVEGP 347		
Db	1768	KCFSDRMTNFNNSIPEWGLSDPQCYDFVATCAQLLLTAFDPVMYQOCWTI-----1819		
Qy	348	YVAFAGRGVVDIRHPYDDPTBPSYNNK-----FLAKOSV 381		
Db	1820	-----PYNCTTPPYGETWTGYNESDFLNGPCYDDSAMEGYLNRPV 1865		
Qy	382	MDAIGV--NINYTQSNNDVYVAFQQTGDPWFNFIEDLEEILA---LPVRVSLIYGADY 436		
Db	1866	RXALNIPDSVPVWAANNIINAYNQVDSIVN---LQIINANPAPNFKMILYSGDAT 1921		
Qy	437	ICNWFGGQ-----AVSLAANYGQAA---QFRSAGYTPLVKNGVEYGETREYGNFSPT 485		
Db	1922	MVNWLGAEIFTANNFAALGLTSSPSRAQWTFQYNST-FQP-TVAGYQTSYTSNAINIDVL 1979		
Qy	486	RVYEAGHEVPYQPIASQLFNRTIFGMDIAGQKKIWPYSKYNTGTATATHTQSSVP--L 543		
Db	1980	TVKSGSHFVPLDRPQALQM-----LYNFVKRGYNTPFDLNSNFNTTTTSTTTTTPGTG 2035		
Qy	544	PTATS 548		
Db	2036	PTVTA 2040		

RESULT 14
A29412

carboxypeptidase C (EC 3.4.16.5) precursor - wheat
M/Alternate names: carboxypeptidase Y homolog
C/Species: Triticum aestivum (common wheat)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999
C/Accession: A29412
J./Biol. Chem. 262, 13726-13735, 1987
A/Title: A gibberellin responsive wheat gene has
A/Reference number: A29412; MUID:98007603; PMID:28

Search completed: November 21, 2003, 18:01:33
Job time : 22 secs

A;Accession: A29412
A;Molecule type: mRNA
A;Residues: 1-499 <BAU>
C;Genetics:
A;Introns: 100/2; 156/3; 196/3; 244/3; 295/1; 324/3; 411/1; 455/3
C;Superfamily: serine carboxypeptidase
C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase; zymogen
F;1-27/Domain: signal sequence #status predicted <SIG>
F;23-73/Domain: propeptide #status predicted <PRO>
F;74-484/Product: carboxypeptidase C #status predicted <MAT>
F;485-499/Domain: carboxyl-terminal propeptide #status predicted <CPRO>
F;144/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;216,404,461/Active site: Ser, Asp, His #status predicted

Query Match 14.0%; Score 416.5; DB 1; Length 499;
Best Local Similarity 24.6%; Pred. No. 4,4e-21;
Matches 131; Conservative 85; Mismatches 231; Indels 85; Gaps 16;

QY 7 LSVLPLVAASWALPGSTPASVGRRLPKNPTG-----VKLTITANNVTTRYKEPGA 58
DB :
1 MATTPRLASLLALLCALCAAAGAUFLPPDASPQAERLIARLNALLPGRPRGLGAGAE 60
QY 59 GVCE-----TTPGVKS-----YGY--VDTSPESTHFWFFFEAHNPETAITLW 101
DB :
61 DVAPGQLLERVTLPLGLEGVGDGLHHAGYLPNTHDARMYFFFSKGGKED-PVVII 119
QY 102 LNCGSGDSLIGLEELGPCHVNSTDDVINPHSNVEVNLLFLSQPLGVGFYSYSTDVG 161
DB :
120 LTGPGGCSELAVFYENGPFETIANMSLVWNKFWDKIENIIFVDPATGTGFGSYSDRD 179
QY 162 SINPVTVGVENSFPAGVQCRYPTIDATLDTTNLAAEAAWEILQGFLSLPSLSRVQSK 221
DB :
180 TRDEAGV-----SNDLYDFLQVFFKKHEF-----VKN 208
QY 222 DFSJWTSYGHYGPAFNHFVEONRIANGSVNGVLNPSLGIINGIIDEAIQAYYP 281
DB :
209 DFFITGESYAGHYIPAFASRVHGNNK-----NEGTHNLKGFAIGNGLTDPAIQKAYT 263
QY 282 EFANVTYGIKAVNETVNYMKFANQMNGCODLISTCKOTNR-TALADYALCAEATNM 340
DB :
264 DYALD-----MLTIQADYDI---NKFIPECEPAIKLCTDGKASCWAAYVVCNSIENSI 316
QY 341 RDNVEGPYYAFAGRGVDIRHPYDDPTPPSYN--RFLAKDSVMDAIGV-NINTQSND 397
DB 317 MKLVCTKNY-----YDVKECEGKLCDVFSNLEKFFGDKAVRQAIQVGDIEFVSCTS 369
QY 398 VYYAFQQQGFVWPNFIE-DLEILLALPVRSLLYGDADYICWFGGAQVSLAANYSOAA 456
DB 370 VYQAMLT-----WMENLEVGIIPALLEGINVLVIYAGEYDLICNWLGNSRWVHSMWSGQK 425
QY 457 QFRSAGYTPLKUNGVEYGETREYGNFSFTRYVEAGHEVPYYPQIASLQLFNR 508
DB 426 DPAKTABSSFLVDDAQAGVLKSHGALSFLKHVNHAGHWPMVDQPKALEMLER 477

RESULT 15
T49079
serine-type carboxypeptidase like protein - Arabidopsis thaliana
N;Alternate names: protein F4F15.110
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
C;Accession: T49079
R;Alcaraz, J.P.; Claubault, G.; Cortet, A.; Mache, R.; Mewes, H.W.; Rudd, S.; Lemcke, K.
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25015
A;Accession: T49079
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-482 <NC>
A;Cross-references: EMBL:AL049711; GSPDB:GN00061; ATSP:F4F15.110
A;Experimental source: cultivar Columbia; BAC clone F4F15
C;Genetics:
A;Gene: ATSP:F4F15.110

[illegible]

Query Match	17.4%;	Score 517.5;	DB 9;	Length 421;
Best Local Similarity	29.5%;	Pred. No. 6e-40;		
Matches 142;	Conservative 70;	Mismatches 185;	Indels 85;	Gaps 17;

QY	51	RYKEPGAEGVCETTPGVKSYSGYDVTSP	---SHTFFWF	FEARHNPTA	ITLWLN	GGPSD	109
Db							
	1	KIKPKIILGI	---DENVTQYGL	DEVEDKH	FFFTFESR	NDPAKDVII	WLN
							57
QY	110	SLIGLFEELGPGCHVNSTPDDY	INPHSMVNE	SNLLFSLQPLG	VGFSGSY	TVDGSIN	PTGV
Db							169
	58	SLTGLFELGPGSSICPOLKPIGN	PYSWMNS	NATVIFLDQ	PVNV	VGFSYS	
							104
QY	170	VENSSEFAGVQGRYPITDATT	DTTNLA	EAAWEILQGL	SLGPLDS	RVQKDS	FWL
Db							229
	105	-----GSSG-----	VSNTVA	AGKVYNELE	FFDQPE	VYVNGK	---DHI
							146

QY 230 YGHHYGAFFNHHFYEQNERIANGSVNGVQLNFNSLGI--NGIIDEAIAQPIYPEFA--- 284
DB 147 YAGHYIVFASEILSHKDR-----NFNLTSVLINGLTDPLTQNYYPEMACGE 195
QY 285 -----VNTYGIKAVNETVNYMKFANQMPNGCODLISTCKQTNETALADYALCAEATNM 339
DB 196 GGEPSVLPSECSAMEDSL-----ERCGLGIESCYDSQ-----SWWSCVPATIIY 239
QY 340 CRDNEVEGYAFAGRGVYDIRHPYDD-----PTPPSYNKKFLAKDSVMDAIGVNINYTQS 394
DB 240 CNAQAALAPYQR-TGRNVYDIRKDCGGNLCYPTLQD-IDDYLNQDYKAEVGAEDVHYES 297
QY 395 -NNDVYAFQGTGFVWPNFIEDLEEILALPVRVSLYIGDADYICNWFPGQAVSLAANY 453
DB 298 CNFDIRNRLFAGDWMKP-YHTAVTDLLNQDLPILVYAGDKDFICNLGNKAWTDVLPWK 356
QY 454 QAAQFRS---AGYPLKXNGVEYGETREYGNFSTRVYEAGHEVYQPIASLQLFNRTI 510
DB 357 YDEEFASQKVRNWTASITDEVA-GEVKSYPKHTYLRVFNCGHVMVFPDVPENALSMVNEWI 415
QY 511 FG 512
DB 416 HG 417

RESULT 3
US-10-084-018-3
; Sequence 3, Application US/10084018
; Publication No. US20020160499A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hawkins, Phillip R.
; Hillman, Jennifer L.
; Lal, Preeti
; Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN SERINE
; CARBOXYPEPTIDASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/10/084,018
; APPLICATION NUMBER: US/10/084,018
; FILING DATE: 25-Feb-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,689A
; FILING DATE: <unknown>
; APPLICATION NUMBER: 08/828,488
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0241 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: MMLR3DT01
; CLONE: 566993
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-084-018-3

Query Match 12.1%; Score 361; DB 14; Length 476;
Best Local Similarity 26.7%; Pred. No. 4.1e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

QY 65 PG--VKYSYGY--VDTSPESHFTFFFEARHNPTAPITLWNLGPGSDSLGLFEEGLP 120
DB 67 PGLNKKSYAGPLTYNKTNSNLFFFPFPAIQPEDAPVLLWLGQPGGSSMXGLFEVHG 126

QY 230 YGHHYGAFFNHHFYEQNERIANGSVNGVQLNFNSLGI--NGIIDEAIAQPIYPEFA--- 284
DB 147 YAGHYIVFASEILSHKDR-----NFNLTSVLINGLTDPLTQNYYPEMACGE 195
QY 285 -----VNTYGIKAVNETVNYMKFANQMPNGCODLISTCKQTNETALADYALCAEATNM 339
DB 196 GGEPSVLPSECSAMEDSL-----ERCGLGIESCYDSQ-----SWWSCVPATIIY 239
QY 340 CRDNEVEGYAFAGRGVYDIRHPYDD-----PTPPSYNKKFLAKDSVMDAIGVNINYTQS 394
DB 240 CNAQAALAPYQR-TGRNVYDIRKDCGGNLCYPTLQD-IDDYLNQDYKAEVGAEDVHYES 297
QY 395 -NNDVYAFQGTGFVWPNFIEDLEEILALPVRVSLYIGDADYICNWFPGQAVSLAANY 453
DB 298 CNFDIRNRLFAGDWMKP-YHTAVTDLLNQDLPILVYAGDKDFICNLGNKAWTDVLPWK 356
QY 454 QAAQFRS---AGYPLKXNGVEYGETREYGNFSTRVYEAGHEVYQPIASLQLFNRTI 510
DB 357 YDEEFASQKVRNWTASITDEVA-GEVKSYPKHTYLRVFNCGHVMVFPDVPENALSMVNEWI 415
QY 511 FG 512
DB 416 HG 417

RESULT 2
US-09-901-252-15
; Sequence 15, Application US/09901252
; Patent No. US20020026658A1
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clint
; TITLE OF INVENTION: Genes Encoding Sinapoylglucose:Malate Sinapoyltransferase and Met
; FILE REFERENCE: N1422-004
; CURRENT APPLICATION NUMBER: US/09/901,252
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216593
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Saccharomyces
US-09-901-252-15

Query Match 17.4%; Score 517.5; DB 9; Length 421;
Best Local Similarity 29.5%; Pred. No. 6e-40;
Matches 142; Conservative 70; Mismatches 185; Indels 85; Gaps 17;

QY 51 RYKEPAGEGVCETTPGVKSYSGYVDTSP-SHTFFFWFFFEARHNPTAPITLWNLGPGSD 109
DB 1 KIKDPKILGI---DPNVTQYTVGLDVEDEKHHFFFTFESRNPDPKDPVILWNLGPGCS 57
QY 110 SLGLFEEELGFCPHVNSTFDDYINPHSNVNSNLLFTSQPLGVGFVSBDTVGDSINFTGV 169
DB 58 SLTGLFELGFSISGPDLPKPIGNPYSNWSNATVIFLDQPVNVGFSY-----104
QY 170 VENSSFAGVQGRYPTIDATLIDITNLAAEAWELQGLSLPSLDSRVQSKDFSLWTES 229
DB 105 -----GSSG-----VSNTVAAGKDVNFELEFFQFPFYVNGQ--DEHIAGES 146
QY 230 YGHHYGAFFNHHFYEQNERIANGSVNGVQLNFNSLGI--NGIIDEAIAQPIYPEFA--- 284
DB 147 YAGHYIVFASEILSHKDR-----NFNLTSVLINGLTDPLTQNYYPEMACGE 195
QY 285 -----VNTYGIKAVNETVNYMKFANQMPNGCODLISTCKQTNETALADYALCAEATNM 339
DB 196 GGEPSVLPSECSAMEDSL-----ERCGLGIESCYDSQ-----SWWSCVPATIIY 239
QY 340 CRDNEVEGYAFAGRGVYDIRHPYDD-----PTPPSYNKKFLAKDSVMDAIGVNINYTQS 394
DB 240 CNAQAALAPYQR-TGRNVYDIRKDCGGNLCYPTLQD-IDDYLNQDYKAEVGAEDVHYES 297

QY 121 CHVNSTFDDYINPHSWNEVSNLLFSLQPLGVGFYSVDTVDSINPVTGVWVNSFAGVQG 180
 Db 127 YVWTSNMTLRDRDPFWTTTXXSMLYIDNPVGTGFSFTDDTHG----- 167
 QY 181 RYPTIDATLDTNLAARAAWEILQGLFSLGSLDSRVQSKDPSLWTSYGGHYGPAPFN 240
 Db 168 -----YAVNEDDVARLDYLSALIOFF-----QIFPEYKNDNFYVTGSEYAGKYVPAIAH 215
 QY 241 HFYEONERIANGVNGVQLNFNSLGIINGIIDEALQAPYEPFAVNNYTGKAVNETVYN 300
 Db 216 LIHSLN-----PVREKINLNGAIGDGYSDPSIIGGYAEF-----LYQIGLLDEKQK 265
 QY 301 YMKFANQMPNGCODLISTCKQTN---RTALADYALCAEATN---MCRDNVEG---PYAFAG 353
 Db 266 Y--FQKQ-----CHECIEHIRKQNWFEAFIILDKLLDGLTSDPSYFQNVTCGSNYNFL- 318
 QY 354 RGVVDIRHPYDDTPPSYNNKELAKDSVMDAIGVNNYNTQSNNDVY---YAFQQTGDFVWP 411
 Db 319 -----RCTEPEQDLYYVKFSLPEVRQAIHVG-NQTFNDGTIVEKYLREDTVQSVKP 369
 QY 412 NFIEDLEILALPVRVSLIYGDADI-----CNWFGQAVSLAANYSQAAQ-- 457
 Db 370 WLTEIMNN-----YKVLINYGQLDIIVAAALTERSLMGMDWKSQ-----EYKKAQKV 418
 QY 458 ---FRS-----AGYTPLKXVNGVEYGETREYGNFSTRVYEAGHEVPYQPIASLQLENRTI 510
 Db 419 WKIFKSDSEVAGY-----IRQVGDHFQVIRGGGHILPYDQPLRAFDMINRFI 466
 QY 511 F--GWD 514
 Db 467 YGKGWD 472

RESULT 4

US-09-909-320-164
 ; Sequence 164, Application US/09909320
 ; Patent No. US20020132240A1

GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, A.
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Hurrey, Austin L.
 ; APPLICANT: Hillan, Kenneth, J.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Mather, Jennie P.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William, I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: 10466-14
 ; CURRENT APPLICATION NUMBER: US/09/909,320
 ; CURRENT FILING DATE: 2002-01-04
 ; PRIOR APPLICATION NUMBER: PCT/US00/04414
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: US 60/143,048
 ; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698
 ; PRIOR FILING DATE: 1999-07-26
 ; PRIOR APPLICATION NUMBER: US 60/146,222
 ; PRIOR FILING DATE: 1999-07-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/20594
 ; PRIOR FILING DATE: 1999-09-08
 ; PRIOR APPLICATION NUMBER: PCT/US99/20944
 ; PRIOR FILING DATE: 1999-09-13
 ; PRIOR APPLICATION NUMBER: PCT/US99/21090
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/21547
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/23089
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: PCT/US99/28214
 ; PRIOR FILING DATE: 1999-11-29
 ; PRIOR APPLICATION NUMBER: PCT/US99/28313
 ; PRIOR FILING DATE: 1999-11-30
 ; PRIOR APPLICATION NUMBER: PCT/US99/28564
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/28565
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/30095
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: PCT/US99/30911
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US99/30999
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US00/00219
 ; PRIOR FILING DATE: 2000-01-05
 ; NUMBER OF SEQ ID NOS: 423
 ; SEQ ID NO 164
 ; LENGTH: 476
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-909-320-164

Query Match

Best Local Similarity 12.1%; Score 359; DB 10; Length 476;
 Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

QY 65 PG--VKSYSGY--VDTSPESTHFFWFFFAHNPETAPITLWNGGPGSLSLGLFEELGP 120
 Db 67 PGLNWKSYAGELTVNKTYSNLSLFFWFFFAQIQEPDAPVVLWLTQGGPGSSMFLFVEHGP 126
 QY 121 CHVNSTFDDYINPHSWNEVSNLLFSLQPLGVGFYSVDTVDSINPVTGVWVNSFAGVQG 180
 Db 127 YVWTSNMTLRDRDPFWTTTXXSMLYIDNPVGTGFSFTDDTHG----- 167
 QY 181 RYPTIDATLDTNLAARAAWEILQGLFSLGSLDSRVQSKDPSLWTSYGGHYGPAPFN 240
 Db 168 -----YAVNEDDVARLDYLSALIOFF-----QIFPEYKNDNFYVTGSEYAGKYVPAIAH 215
 QY 241 HFYEONERIANGVNGVQLNFNSLGIINGIIDEALQAPYEPFAVNNYTGKAVNETVYN 300
 Db 216 LIHSLN-----PVREKINLNGAIGDGYSDPSIIGGYAEF-----LYQIGLLDEKQK 265
 QY 301 YMKFANQMPNGCODLISTCKQTN---RTALADYALCAEATN---MCRDNVEG---PYAFAG 353
 Db 266 Y--FQKQ-----CHECIEHIRKQNWFEAFIILDKLLDGLTSDPSYFQNVTCGSNYNFL- 318
 QY 354 RGVVDIRHPYDDTPPSYNNKELAKDSVMDAIGVNNYNTQSNNDVY---YAFQQTGDFVWP 411
 Db 319 -----RCTEPEQDLYYVKFSLPEVRQAIHVG-NQTFNDGTIVEKYLREDTVQSVKP 369
 QY 412 NFIEDLEILALPVRVSLIYGDADI-----CNWFGQAVSLAANYSQAAQ-- 457
 Db 370 WLTEIMNN-----YKVLINYGQLDIIVAAALTERSLMGMDWKSQ-----EYKKAQKV 418
 QY 458 ---FRS-----AGYTPLKXVNGVEYGETREYGNFSTRVYEAGHEVPYQPIASLQLENRTI 510
 Db 419 WKIFKSDSEVAGY-----IRQVGDHFQVIRGGGHILPYDQPLRAFDMINRFI 466

QY 511 F--GWD 514
Db 467 YGKGWD 472

RESULT 5

US-09-909-088B-164

; Sequence 164, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,088B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219

; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-088B-164

Query Match 12.1%; Score 359; DB 10; Length 476;

Best Local Similarity 26.7%; Pred. No. 6.3e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

QY 65 PG--VKSYGY--VDTSPESHTEFFPEARHNPTAPITLWNGGPGSDSLIGLFEELGP 120
Db 67 PGLNMKSYAGFLTVMKTYNSNLFWFPPAQIQPDAPVWLWQGGPGSSMFGFVEHGP 126
QY 121 CHVNSTFDDYINPHSMNEVSNLLEQLSGVGFYSYDTSVDSINPVTGVVSNSSFAGVQG 180
Db 127 YVVTSNMTLRDRDPFWTTLSMLYIDNPVGTFSTDDTHG----- 167
QY 181 RYPTIDATLIDTNLAARAAWEILOGELSLGPLSDRSVQSKDSLWTSYGGHYGPAPFN 240
Db 168 -----YAVNEDDVARDLYSALIQFF-----QIFPEYKKNDFYVTGSAKYVPALAH 215
QY 241 HFYCNERIANGVNGVQLNFNSGLINGIIDEALQAPYEPFAVNNYGIKAVNETVYN 300
Db 216 LIHSLN-----PVREVKINLNGIAGDGYSDPSIIGGYAEF-----LYQIGLLEKQKK 265
QY 301 YMKFANQMPNGCQDLISTCKQTN---RTALADYALCAPATN--MCRDNVEG--PYVAFAG 353
Db 266 Y--FOKQ---CHECIEHKKQWPEAFELDKLDGLTSDPSYFQNVGTGCSYNYFL- 318
QY 354 RGVDIRHPYDDPTPPSYNKFSLAKSDYMDAIGVNNYVTSQNDYV--YAFQOTGDFVWP 411
Db 319 -----RCTEPEDQLYYKFLSLPEVRQAIHVG--NOTFNDGTIVEKYLREDTIVQSVAP 369
QY 412 NFIEDLEELALPVRVSLIYGDADYI-----CNWFGQAVALSANYSOAAQ-- 457
Db 370 WLTEIMNN-----YKVLVYNGQLDIIVAAALTERSIMGMDWKSQ-----EYKKAEEKV 418
QY 458 ---FRS---AGYTLKVNVEYGETREYVGNFSFTRVVEAGHEVVPYQIASLQLFNRI 510
Db 419 WKIFKSDSEVAGY-----IRQAGDFHQVIIRGGHILPYDQPLRAFDMINRFI 466
QY 511 F--GWD 514
Db 467 YGKGWD 472

RESULT 6

US-09-905-291A-164
; Sequence 164, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James

```

; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905/291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-291A-164

Query Match      12.1%; Score 359; DB 10; Length 476;
Best Local Similarity 26.7%; Pred No. 6.3e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

QY 65 PG--VKSYSGY--VDTSPESHFFWFFFAHNPETAPITLWNGGPGSLLGLFEEELGP 120
D 67 PGLANKSVAGFLTKNTYNSLFFWFFPAQIQEPAPVVLWLGPGGSGMFLFVEHGP 126
QY 121 CHVNSTFDYINPHSWNEVSNLLFLSQPLGVGFYSYDVTVDGSIINPVTVGVNSSPAGVQG 180
D 127 YVTSNMTLRDRDFPWTTLTSLMLYIDNPVGTGFSFTDTHG----- 167
QY 181 RYPTIDATLIDTNLAEEAAWEILGGLSGLPSLDSRVQSKDPSLWTSYGHYGPAPFN 240
D 168 -----YAVNEDEVARDLYSALIGFF-----QIPPEYKNDFFVTGSGYAGKYVPALAH 215
QY 241 HFYQNERIANGSVNGVQNFNSLGIINGIIDEAQAPYEPFAVNNYTGKAVNETVYN 300
D 216 LIHSLN-----PVREVKINLNGIAGDYSPIGIGYAEF-----LYQIGLDEKQKK 265
QY 301 YMKFANQMPNCQDLISTCKQTN-----RTALADYALCABATN--MCRDNVEG--PYAPAG 353

```

```

Db 266 Y--FOKQ---CHCIEHIRKQNWFEAFILDKLIDGLTSDPSYFQVNTGCSNYNFI-- 318
QY 354 RGVYDIRHPYDDPPPSYKFLAKDSYMDAIGVNNINVTQSNNDVY--YAFQOTGDFWVP 411
D 319 -----RCTEPEDQLYYVAFSLPEVRQAIHVG--NOTFNDGTIVEKYLRDITVQSVKP 369
QY 412 NFIEDLEILALPVRVSLIYGDADYI-----CNWFGGQAVSLAANYSOAAQ-- 457
D 370 WLTEIMNN-----YKVLIIYNGQLDIIIVAAALTEKSLMCMQWKGQ-- 418
QY 458 ---FRS---AGYTPLKVGVEYGETREYGNFSFTRVFEAGHEVYVYQPIASLOLFNRTI 510
D 419 WKIFKSDSEVAGY-----IRQAGDFHQVIRGGGHILPYDQPLRAFDMINRFI 466
QY 511 F--GWD 514
D 467 YGKGWD 472

```

RESULT 7

```

US-09-902-853-164
; Sequence 164, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,853
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US/09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05

```

PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 164
LENGTH: 476
TYPE: PRT
ORGANISM: Homo Sapien
US-09-902-853-164

Query Match 12.1%; Score 359; DB 10; Length 476;
Best Local Similarity 26.7%; Pred. No. 6.3e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;
QY 65 PG--VKSXGY--VDTSPESHFFWFFFAHNPETAPITLWNGGPGSDSLGLPEELGP 120
DB 67 PGLNKKVAGELTVNKTNSNLFPPFFPAQIQPEAPVVLWQGGPGSSMFLVHEGP 126
QY 121 CHVNSTEDYINPHSNEVSNLLFSLQPLGVGSFSDTVGSGINPVTVGVVNSFAGVQG 180
DB 127 YVTSNMLURDPFWTTLSMLYIDNPVGTGFSFTDTHG----- 167
QY 181 RYPTIDATLIDTTNLAAEAAWEILOGLSLPSLSRVSKOFSLWTSYGHYGPAPFN 240
DB 168 -----YAVNEDDVARDLYSALIQF-----QIFPEYKNDFFVTGESYAGKYVPAIAH 215
QY 241 HFEQNERANGSVNGVQLNFSNLGIINGIIDEAIOAPYYPFAVNTYGIKAVNNTVYN 300
DB 216 LHSNL-----PVREKINLNGIAGDGSDESIIIGYAEF-----LTOIGLLDEKQKK 265
QY 301 YMKFANQMPGCDLSTCKQTN---RTALADYALCAEATN---MCRDNVEG---PYAFAG 353
DB 266 Y--FQKQ-----CHECHIRKQWFEAFBLDKLDGDLTSDPSYFQNVGTGCSNYNFL- 318
QY 354 RGVYDIRHPYDDPTPPSYNKLAKOSVMDAIGVNNINYTQSNNDVY--YAFQOTGDFVWP 411
DB 319 -----RCTEPEDQLYVVKFLSLPEVRQAIHVQ-NQTFNDGTIVEKYLREDTVSQVKP 369
QY 412 NTFIDLEEILALPVRSLLIYGDADYI-----CNWFGGQAVSLAANTSOAAQ-- 457
DB 370 WLTEIMNN---YKVLTYNGQDIIIVAAALTELSLGMGMWKGSO-----EYKAEKV 418
QY 458 ---FRS-----AGYTPLKXVNGVEYGETREYGNFSTRVYEAGHEVPYQPIASIQLFNRTI 510
DB 419 WKLFKSDSEVAGY-----IFQAGDFHQVIRGGCHILFYDQPLRAFDMINRFI 466
QY 511 F--GWD 514
DB 467 YGKGWD 472

RESULT 8
US-09-907-824-164
Sequence 164, Application US/09907824
Publication No. US20020197671A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc

APPLICANT: Baton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Macher, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/907,824
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: 09/665,350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 164
LENGTH: 476
TYPE: PRT
ORGANISM: Homo Sapien
US-09-907-824-164

Query Match 12.1%; Score 359; DB 10; Length 476;
Best Local Similarity 26.7%; Pred. No. 6.3e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

Query Match	12.1%; Score 359; DB 11; Length 476;
Best Local Similarity	26.7%; Pred. No. 6,3e-25;
Matches 130; Conservative	62; Mismatches 178; Indels 116; Gaps 21;
QY	65 PG--VXSYSY--VDTSPESHTEFFFEARHNHPETAPITLWLNKGGPSDSLGLFELGIP 120
DB	67 PGLNMSYAGFLVNTKYNLFFWPPFAIQPEDAPWLWLGQPGSGSMFLGVEHGP 126
QY	121 CHVNSTEDDYNHSMNEVSNLLFLSQPLGVGHSYSDTVGDSINPTVGVVNSSFAGVQG 180
DB	127 YVVTNNMTLRDRDPPTTILSMLYIDNPVGTFSTDDTHG----- 167
QY	181 RYPTIDATLDTNLAEEAAWEILQGLSLGSLDSRVQSKDFSLWTSYGGHYGPAFFN 240
DB	168 -----YAVNEDDVARDLYSALIQFF-----QIFPEYKKNDFVVTGESYAGKYVPAIAH 215
QY	241 HFYEQNERIANGSVNGQLNFSNGIINGIIDEAIAQIAPYEPFPAVNNTYGIKAVNETVYN 300
DB	216 LIHSLN-----PVREVKINLNGIAGDYSDBPSIIGGYAEF-----LYQIGLLDEKQKK 265
QY	301 YMKFANQMPNGCQDLISTCKQTN---RTALADYALCAEATN--MCRDNVEG--PYAFAG 353
DB	266 Y--FQKQ---CHECIEHIRKQNWFEAFELDKLLDGLTSDPSYFQVNTGCSNYNPL- 318
QY	354 RGVDYDRHPYDDPTPPSYNKFLLAKDSVMDAIGVNINVTQSNNDVY--YAFQQTGDFVYP 411
DB	319 -----RCEPEDQLYVVKFLSLPEVRQAIHVG--NOTFNDGTIVKYLREDTVQSVKP 369
QY	412 NFIEDLEILALPVRSLLIYGADYI-----CNMFGGQAVSLAANYSOAAQ-- 457
DB	370 WITEIMNN-----YKVLINGQLDIIIVAAALTERSLMGMDWKGSQ-----EYKKAERKV 418
QY	458 ---FRS-----AGYTPLKNGVEYGETREYGNFSTRVYEAGHEVPYQIPLASQLQFNRTI 510
DB	419 WKIFKSDSEVAGY-----IROAGDFHQVIIRGGGHILEFDQPLAFDMINRPI 466
QY	511 P--GWD 514
DB	467 YGKGWD 472
RESULT 11	
US-09-906-742-164	
; Sequence 164, Application US/09906742	
; Publication No. US20030023054A1	
; GENERAL INFORMATION:	
; APPLICANT: Genentech, Inc.	
; APPLICANT: Ashkenazi, Avi	
; APPLICANT: Botstein, David	
; APPLICANT: Desnoyers, Luc	
; APPLICANT: Eaton, Dan L.	
; APPLICANT: Ferrara, Napoleone	
; APPLICANT: Filvaroff, Ellen	
; APPLICANT: Fong, Sherman	
; APPLICANT: Gao, Wei-Qiang	
; APPLICANT: Gerber, Hanspeter	
; APPLICANT: Gerritsen, Mary E.	
; APPLICANT: Goddard, A.	
; APPLICANT: Goddowski, Paul J.	
; APPLICANT: Grimaldi, Christopher J.	
; APPLICANT: Gurney, Austin L.	
; APPLICANT: Hillan, Kenneth, J.	
; APPLICANT: Kijavini, Ivar J.	
; APPLICANT: Mather, Jennie P.	
; APPLICANT: Pan, James	
; APPLICANT: Paoni, Nicholas F.	
; APPLICANT: Roy, Margaret Ann	
; APPLICANT: Stewart, Timothy A.	
; APPLICANT: Tumas, Daniel	
; APPLICANT: Williams, P. Mickey	
; APPLICANT: Wood, William, I.	
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic	
; TITLE OF INVENTION: Acids Encoding the Same	
; FILE REFERENCE: 10466-14	
; CURRENT APPLICATION NUMBER: US/09/904,011	
; CURRENT FILING DATE: 2001-07-11	
; PRIOR APPLICATION NUMBER: 09/665,350	
; PRIOR FILING DATE: 2000-09-18	
; PRIOR APPLICATION NUMBER: PCT/US00/04414	
; PRIOR FILING DATE: 2000-02-22	
; PRIOR APPLICATION NUMBER: US 60/143,048	
; PRIOR FILING DATE: 1999-07-07	
; PRIOR APPLICATION NUMBER: US 60/145,698	
; PRIOR FILING DATE: 1999-07-26	
; PRIOR APPLICATION NUMBER: US 60/146,222	
; PRIOR FILING DATE: 1999-07-28	
; PRIOR APPLICATION NUMBER: PCT/US99/21547	
; PRIOR FILING DATE: 1999-09-15	
; PRIOR APPLICATION NUMBER: PCT/US99/23089	
; PRIOR FILING DATE: 1999-10-05	
; PRIOR APPLICATION NUMBER: PCT/US99/28214	
; PRIOR FILING DATE: 1999-11-29	
; PRIOR APPLICATION NUMBER: PCT/US99/28313	
; PRIOR FILING DATE: 1999-11-30	
; PRIOR APPLICATION NUMBER: PCT/US99/28564	
; PRIOR FILING DATE: 1999-12-02	
; PRIOR APPLICATION NUMBER: PCT/US99/28565	
; PRIOR FILING DATE: 1999-12-02	
; PRIOR APPLICATION NUMBER: PCT/US99/30095	
; PRIOR FILING DATE: 1999-12-16	
; PRIOR APPLICATION NUMBER: PCT/US99/30911	
; PRIOR FILING DATE: 1999-12-20	
; PRIOR APPLICATION NUMBER: PCT/US99/30999	
; PRIOR FILING DATE: 1999-12-20	
; PRIOR APPLICATION NUMBER: PCT/US00/00219	
; PRIOR FILING DATE: 2000-01-05	
; NUMBER OF SEQ ID NOS: 423	
; SEQ ID NO 164	
; LENGTH: 476	
; TYPE: PRT	
; ORGANISM: Homo Sapien	
; US-09-904-011-164	

```
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/906,742
CURRENT FILING DATE: 2001-07-16
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 164
LENGTH: 476
TYPE: PRT
ORGANISM: Homo Sapien
US-09-906-742-164

Query Match
Best Local Similarity 12.1%; Score 359; DB 11; Length 476;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

65 PG--VKSYSGY--VDTSPESHTEFFFEARRHNPETAPITLWNGPGSDSLGLFEEELGP 120
67 PGLANKSYAGELTVNKTYSNLFWFFFPAQIQEPDAPVVLWLGQPGGSSMGLFVEHGP 126
121 CHVNSTEDYINPHSNVSNLFLSOLPLGVGFSYSDTVGSGINPVTGVVENSFAGVQG 180
127 YVTSNMTLRDPPFTMTLSMLYDINPVTGFSFTDTHG-----167
181 RYPTIDAFTIDTTLAAAEWELIGLGLSPLDSRVQSKDFSLWTSYGHYGPAPFN 240
168 -----YAVNEDDVARDLYSALIQFF-----QIPPEYKNDFFYTGSGYAGKYVPAIAH 215
241 HFYQNEIRANGSVNGVQNFNSLGIINGIIDEALQAPYEPFAVNNYTGKAVNETVYN 300
216 LIHSIN-----PVREVKNLNGIAIGDYSPPESIIIGYAEF-----LYQIGLDEKQKK 265
301 YMKFANQPNQCQDLISTCKQTN---RTALADYALCAEATN---MCRDNVEG---PYAPAG 353
266 Y--FQKQ---CHECIEHQRQNWFPEAFIELDKLLDGLTSDPSYFQNVTCSSNYNPL- 318
354 RGVYDIRHPYDPPSPSYNKFELAKDSYMDAIGNVNTQSNNDV---YAFQQTGDFWVP 411
319 -----RCTEPEDQLYVFKLSLPEVFRQAIHVG-NQTFNDGTIVKYLREDTVQSVKP 369

FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/906,742
CURRENT FILING DATE: 2001-07-16
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 164
LENGTH: 476
TYPE: PRT
ORGANISM: Homo Sapien
US-09-906-742-164

US-09-906-838-164
Sequence 164, Application US/09906838
Publication No. US20030027143A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/906,838
CURRENT FILING DATE: 2001-07-16
PRIOR APPLICATION NUMBER: 09/665,350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 164
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-906-838-164

Query Match      12.1%; Score 359; DB 11; Length 476;
Best Local Similarity 26.7%; Pred. No. 6.3e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

Qy 65 PG--VKSYGY--VDTSPESHTEFFFEARHNPTAPITLWNGGPGSDSLIGLFEELGP 120
Db 67 PGLNWKSYAGFLTVNKTNSNLFWFPPAQIQEDAPVVLWLOGGPGSSMFGFLVBEHGP 126
Qy 121 CHVNSTFDYINPHSWNEVSNLLFLSQPLGVGFSDYDVGSDINPVTGVWENSSPAGVQG 180
Db 127 YVVTSMNLTDRDPFWITTLNLAEEAWEILQGLSLDSRVQSKDFSLWTSYGGHYGPAPFN 167
Qy 181 RYPTIDATLIDTNLAEEAWEILQGLSLDSRVQSKDFSLWTSYGGHYGPAPFN 240
Db 168 -----YAVNEDDVARDLYSALIOFF-----QIPPEYKNNDFYVTGSGYACKYVPAIAH 215
Qy 241 HFEQNERIANGSVNGVQNFNSLGIINGIDEAQAIPYEPFVANNYTGKAVNETVYN 300
Db 216 LIHSLN-----PVREVKINLNGIAGDGYSDPESIIIGYAEF-----LYQIGLLDEKQKK 265
Qy 301 YMKPANOMPGQDILSTCKQTN---RTALADYALCAEATN---MCRDNVEG--PYAFAG 353
Db 266 Y--FOKQ----CHECHIEHTRKQNFWEFAFEILDKLGDITSDPSFQNTGSGNYNLF- 318
Qy 354 RGVDYIRHPDPTPSYKFKIADKSVMDAIGVNIYNTQSNNDVY--YAFQGTGFWVP 411
Db 319 -----RCTSPEDQLYYVKFLSLPEVQAIVHG-NQTFNDGTIVEKYLREDTVQSVKP 369
Qy 412 NFIEDLEEILALPVRVSLYGDADYI-----CNWFGGQAVSLAANYSQAAQ-- 457
Db 370 WLTEIMNN-----YKVLIVNGQLDIIVAAALTERSLMGMDWKGSQ-----EYKKAERKV 418
Qy 458 ---FRS---AGVTPKLVNGVEYGETREYGNFSTFVWZBAGHEVPPYQPIASIQLPNRTI 510
Db 419 WKIFKSDSEVAGY-----IRQAGDFHQVIIRGGHILPYDQPLRAEDMINRFI 466
Qy 511 F--GWD 514
Db 467 YGKGWD 472

RESULT 13
US-09-907-613-164
; Sequence 164, Application US/09907613
; Publication No. US20030027145A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
```

```

Db 127 YVVTSMNLTDRDRFPWTTLTSLMLYIDNPVGTGFSFTDDTHG----- 167
QY 181 RYPTIDATLIDTTNLAAEAAWEILQGLFSLDSRVQSKDFSLWTSYGHYGPAPFN 240
Db 168 -----YAVNEDDVARLDYLSALIOFF-----QIFPEYKNDFFVTGESYAGKYVPAIAH 215
QY 241 HFYQNERIANGSVNGVQLNFSGLIINGIIDEAIAQAPYPEFAVNNTYGIKAVNETVYN 300
Db 216 LIHSLN-----PVEVKINLNGIAIGDYSDPESIIGGYAEF-----LYQIGLLDERQKK 265
QY 301 YMKFANQMPNGCQDLISTCKQTN---RTALADYALCABATN---MCRDNVEG---PYAFAG 353
Db 266 Y--FQKQ---CHECIEHQRKNWFEAFEILDKLDDGLTSDPSYFQNVTCGSNYNFI-- 318
QY 354 RGVYDIRHPYDDPTPPSYNKFELAKDSVMDAIGVNIYNTQSNNDVY--YAFQQTGDFWVP 411
Db 319 -----RCTEPEDQLYVVKFSLPEVRQAIHVG--NOTFNDGTIVEKYREDTIVQSKVP 369
QY 412 NFIEDLEEILALPVRVSLIYGDADYI-----CNWFGGQAVSLAANYSQAAQ-- 457
Db 370 WLTEIMNN-----YKVLINQGLDIIVAALATERSLMGMWKGSGQ-----EYKKAEEKV 418
QY 458 ---FRS---AGYTPLKVGVEYGETREYGNFSTRVYEAQHEVPYQYQPIASLOLFNRTI 510
Db 419 WKIFKSDSEVAGY-----IRQAGDFHOVIIRGGGHILPYDQPLRAFDMINRFI 466
QY 511 F--GWD 514
Db 467 YGKGWD 472

```

RESULT 14

```

US-09-907-942-164
; Sequence 164, Application US/09907942
; Publication No. US20030027146A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,942
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222

```

```

Query Match 12.1%; Score 359; DB 11; Length 476;
Best Local Similarity 26.7%; Pred. No. 6.3e-25;
Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;

QY 65 PG--VKSYSGY--VDTPESHTFFWFPEARHPETAPITLWLNCGPGSDSLGLFELGP 120
Db 67 PGLNMKSYAGFLTVKNTYNSNLFNFFFPQIQPEDAPVVLWQGGPGGSMFLFVEHGE 126
QY 121 CHVNSTFDDYINPHSWNEVSNLFLSQPLGVGFYSYSDTVDGSIINPVTGVVENSFAGVQ 180
Db 127 YVVTSMNLTDRDRFPWTTLTSLMLYIDNPVGTGFSFTDDTHG----- 167
QY 181 RYPTIDATLIDTTNLAAEAAWEILQGLFSLDSRVQSKDFSLWTSYGHYGPAPFN 240
Db 168 -----YAVNEDDVARLDYLSALIOFF-----QIFPEYKNDFFVTGESYAGKYVPAIAH 215
QY 241 HFYQNERIANGSVNGVQLNFSGLIINGIIDEAIAQAPYPEFAVNNTYGIKAVNETVYN 300
Db 216 LIHSLN-----PVEVKINLNGIAIGDYSDPESIIGGYAEF-----LYQIGLLDERQKK 265
QY 301 YMKFANQMPNGCQDLISTCKQTN---RTALADYALCABATN---MCRDNVEG---PYAFAG 353
Db 266 Y--FQKQ---CHECIEHQRKNWFEAFEILDKLDDGLTSDPSYFQNVTCGSNYNFI-- 318
QY 354 RGVYDIRHPYDDPTPPSYNKFELAKDSVMDAIGVNIYNTQSNNDVY--YAFQQTGDFWVP 411
Db 319 -----RCTEPEDQLYVVKFSLPEVRQAIHVG--NOTFNDGTIVEKYREDTIVQSKVP 369
QY 412 NFIEDLEEILALPVRVSLIYGDADYI-----CNWFGGQAVSLAANYSQAAQ-- 457
Db 370 WLTEIMNN-----YKVLINQGLDIIVAALATERSLMGMWKGSGQ-----EYKKAEEKV 418
QY 458 ---FRS---AGYTPLKVGVEYGETREYGNFSTRVYEAQHEVPYQYQPIASLOLFNRTI 510
Db 419 WKIFKSDSEVAGY-----IRQAGDFHOVIIRGGGHILPYDQPLRAFDMINRFI 466
QY 511 F--GWD 514
Db 467 YGKGWD 472

```

Matches 130; Conservative 62; Mismatches 178; Indels 116; Gaps 21;
QY PG--VKSYSY--VDTSPESHTRFWRPPEARHNPETAPITLWNGPGQSDGLIGLFEELGP 120
Db PGLNMSYAGFLVTKNTYNSNLFVFPFPAQLQPEDAPVVLWLOGPGGSSMFLFVEHGP 126
QY 121 CHVNSTEDDYINPHSMNEVSNLLFLSQPLGVGFYSYDTPDGSINPVTGVVENSFFAGVQG 180
Db 127 YVVTSNMTRDRDPFWTTLSMLYIDNFGVGFSTDDTHG----- 167
QY 181 RYPTIDATLIDTNNLAABAAWEILQGLFLPSLDSRVQSKDFSLWTESGHHYGPAPFN 240
Db 168 -----YAVNEDDVARLYSALIQFF-----QIFPEYKNNDFVVTGESYAGKVPALAH 215
QY 241 HFYEQNERLANGSVNGVQLNFNSLGIINGIIDEALQAPYYPFEFAVNNYTGKAVNETVYN 300
Db 216 LIHSLN-----PVREVKINLNGIAGDGYSDPESITGGYAEF-----LYQIGLDEKQK 265
QY 301 YMKFANQMPNGCQDLISTCKQTN---RTALADYALCAEATN--MCRDNVEG--PYIYAFAG 353
Db 266 Y--FQKQ---CHCEIHIRKQNWFEAFELDKLDDGLTSDPSYFQNVTCNSYNYFL- 318
QY 354 RGVDYIRHPYDDPTPPSYNKKELAKDSVMDAIGVNNVYQSNNDYV--YAFQQTGDFYWP 411
Db 319 -----RCTEPEQLYYVRFSLPEVROAIHVG-NQTFNDGTIVEKYLREDITVQSVKP 369
QY 412 NFIEDLEILALPVRVSLIYGDADVI-----CNWFGGOAVSLAANYSOAAQ-- 457
Db 370 WLTEIMNN-----YKVLINYGQDITIIAAALTERSLMGMDWKGSO-----EYKKAEEKV 418
QY 458 ---FRS---AGYTPKLVNGVEYGETREYGNFSTRVYEAGHEVPYQDIASLQLENRTI 510
Db 419 WKIFKSDSEVAGY-----IRQAGDPHQVILGGGHHILPYDQPLRAFDMINRTI 466
QY 511 F--GWD 514
Db 467 YGKGWD 472

Search completed: November 21, 2003, 18:02:55
Job time : 37 secs

RESULT 15
US-09-796-753-40
; Sequence 40, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 40
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-753-40

Query Match 12.1%; Score 359; DB 11; Length 476;
Best Local Similarity 26.7%; Pred. No. 6.3e-25;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 17:58:46 ; Search time 21 Seconds
(without alignments)
1118.216 Million cell updates/sec

Title: US-09-712-338-2
Perfect score: 2979
Sequence: 1 MRGYEFUSVLPVLAASWALP.....HTQSSVPLPTATSMSSVGMA 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCUTS COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2975	99.9	554	3	US-08-943-714-2
2	894	30.0	423	3	US-08-943-714-9
3	630	21.1	557	1	US-08-309-341-2
4	630	21.1	557	1	US-08-608-267-2
5	630	21.1	557	1	US-08-608-452-2
6	630	21.1	557	1	US-08-608-224-2
7	630	21.1	557	2	US-08-967-149-2
8	623	20.9	557	1	US-08-309-341-4
9	623	20.9	557	1	US-08-608-267-4
10	623	20.9	557	1	US-08-608-452-4
11	623	20.9	557	1	US-08-608-224-4
12	623	20.9	557	2	US-08-967-149-4
13	568.5	19.1	491	1	US-09-640-305-4
14	568.5	19.1	491	1	US-08-360-673-4
15	523.5	17.6	532	2	US-08-893-324-33
16	523.5	17.6	532	3	US-08-329-892B-33
17	513.5	17.2	421	2	US-08-807-263-4
18	361	12.1	476	2	US-08-828-488-3
19	361	12.1	476	4	US-09-299-689A-3
20	348.5	11.7	477	2	US-08-828-488-1
21	348.5	11.7	477	4	US-09-299-689A-1
22	336	11.3	480	2	US-08-828-488-8
23	336	11.3	480	4	US-09-299-689A-8
24	336	11.3	480	4	US-09-702-705-336
25	336	11.3	480	4	US-09-736-457-336
26	312.5	10.5	471	2	US-08-828-488-7
27	312.5	10.5	471	4	US-09-299-689A-7

28	267	9.0	523	3	US-08-943-714-11	Sequence 11, Appl
29	246.5	8.3	446	1	US-08-665-966-10	Sequence 10, Appl
30	246.5	8.3	446	3	US-09-041-780-10	Sequence 10, Appl
31	208	7.0	530	3	US-08-943-714-12	Sequence 12, Appl
32	204.5	6.9	481	3	US-08-943-714-10	Sequence 10, Appl
33	193	6.5	35	3	US-08-943-714-5	Sequence 5, Appl
34	172	5.8	179	1	US-08-665-966-8	Sequence 8, Appl
35	172	5.8	179	3	US-09-041-780-8	Sequence 8, Appl
36	169.5	5.7	351	2	US-08-828-488-5	Sequence 5, Appl
37	169.5	5.7	351	4	US-09-299-689A-5	Sequence 5, Appl
38	142	4.8	29	3	US-08-943-714-6	Sequence 6, Appl
39	131	4.4	618	3	US-09-199-290-34	Sequence 34, Appl
40	127.5	4.3	771	4	US-09-634-238-229	Sequence 229, App
41	127.5	4.3	771	4	US-09-634-238-262	Sequence 262, App
42	126	4.2	385	2	US-08-387-942C-24	Sequence 24, Appl
43	126	4.2	1403	2	US-08-387-942C-3	Sequence 3, Appl
44	120	4.0	385	2	US-08-387-942C-27	Sequence 27, Appl
45	120	4.0	872	2	US-08-387-942C-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1

US-08-943-714-2
; Sequence 2, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berk, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dambmann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,714
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4990.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 554 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-943-714-2

Query Match

99.9% Score 2975; DB 3; Length 554;

```

FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-943-714-9

Query Match 30.0%; Score 894; DB 3; Length 423;
Best Local Similarity 40.1%; Pred.No.3.5e-73;
Matches 186; Conservative 66; Mismatches 160; Indels 52; Gaps 8;

QY 59 GVCETTPGVKSGVYDTPSPGHTFPWFPEARHNPTAITLWLGPGSDSLIGLPEEL 118
DB 6 GICETTPGVNQYSGVLSGSNNVMWFFPEARNPQOAPLAAPNGFGCGSSMIGLFQEN 65
QY 119 GPCH-VNSTFDDYINPHSNWEYSNLLFLSPLCGVGSYSDDTVGDSINPVTGVVENSFAG 177
DB 66 GPCHFVNGDSTSLNENSNVYANMIYIDQPIGVGFSYG----- 104
QY 178 VQGRVPTIDATLIDITNLAAEAAWEILQFLSGLPSLDSRVQSKDFSLWTESYGGHYGPA 237
DB 105 -----TDDVT---STVTAAPYVNNLLQAFYACRPYESR----DFAIFTESYGGHYGPE 151
QY 238 FNNHFVEQNERIANGSVNGVQLNFSIIGLIIIDEAQAPYPPEFAVNNTYGKAVNET 297
DB 152 FASYIEQNAAIKAGSVTQGVNVALGVNNGWIDSTIOEKAYIDFSYNNSYQIIDSST 211
QY 298 VVNMKFAQMPCGQDLISTCKQTNRITALADYALCAEATNMCRDNVEGPPYAFAGRGVY 357
DB 212 RDSLLDAYN---NQCLPALQCCSQSGT-----SDCTNADSVCYQIEGPISSSGDFDVY 263
QY 358 DIRHPYDDTPPPSYNFKLAKDSVMDAIGVNNYTSQNNVDVYAFQCTGDFWPNFIEDL 417
DB 264 DIREPSNDPPPKTYSTYLSDPYVKAIGARTNYQECNPGPNKFASTGDNP-RSFLSTL 322
QY 418 EEITALPVRYSLIVGDADYICNVEGGQAVSLAANYSQAAQPSRAGYTPFLKVNGBYGETR 477
DB 323 SSVVQSGNINVLWAGDADWICNLGNTEVANADVPPGNAQFSALDLAPYTVNGVKGQPK 382
QY 478 EYGNFSTRVYEAGHEVPYIOPIASLQLFNRTIFGWDIAEQKK 521
DB 383 TVDNFSLKVVYAGHEVPYIQDTALQAFKOII-----QKK 418

RESULT 3
US-08-309-341-2
; Sequence 2, Application US/08309341
; Patent No. 5594119
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: No. 5594119o No. 5594119disk of No. 5594119ch America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

Best Local Similarity 100.0%; Pred. No. 4.4e-263;
Matches 554; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1  MRGYEFLSVLPVLAASWALPGSTPASVGRRLPKNPTGVKTLTTANNVTIRKEPGAEGV 60
|
|
|
1  MRGYEFLSVLPVLAASWALPGSTPASVGRRLPKNPTGVKTLTTANNVTIRKEPGAEGV 60
|
|
|
61  CETTPGVKSYGYVDTPSHTFFWFFFEARHNPETAITLWNGGPGSDSLTGLPEELGP 120
|
|
|
61  CETTPGVKSYGYVDTPSHTFFWFFFEARHNPETAITLWNGGPGSDSLTGLPEELGP 120
|
|
|
121  CHVNSTFDDYINPHSWNEVSNLLFSLQPLGVGFYSXSDTVDGSINPVTGVVENSFPAGVQG 180
|
|
|
121  CHVNSTFDDYINPHSWNEVSNLLFSLQPLGVGFYSXSDTVDGSINPVTGVVENSFPAGVQG 180
|
|
|
181  RYPTTIDATLIDTNNLAEEAAWEILOGLSLPSLDSRVQSKDFSLWSTESYGGHYGPAFFN 240
|
|
|
181  RYPTTIDATLIDTNNLAEEAAWEILOGLSLPSLDSRVQSKDFSLWSTESYGGHYGPAFFN 240
|
|
|
241  HYEQNERIANGSVNGVQLNFNSLGIINGIIDAIQAPYYPEFAVNNTYGIKAVNETVYN 300
|
|
|
241  HYEQNERIANGSVNGVQLNFNSLGIINGIIDAIQAPYYPEFAVNNTYGIKAVNETVYN 300
|
|
|
301  YMKFANQMPGCGDILSTCKQTNRITALADVALCAEATNMCRDNVEGPIYAFAGRGVYDIR 360
|
|
|
301  YMKFANQMPGCGDILSTCKQTNRITALADVALCAEATNMCRDNVEGPIYAFAGRGVYDIR 360
|
|
|
361  HPYDDTPPSYKNKFLAKOSVMDAIGVINVTQSNNDVYAFQOTGDFVWPNFIEDLEEI 420
|
|
|
361  HPYDDTPPSYKNKFLAKOSVMDAIGVINVTQSNNDVYAFQOTGDFVWPNFIEDLEEI 420
|
|
|
421  LALPVRVSLITYGDADYICWNFGGQAVSLAANYGQAQFRSAGYTPLVKNGVEYGETREYG 480
|
|
|
421  LALPVRVSLITYGDADYICWNFGGQAVSLAANYGQAQFRSAGYTPLVKNGVEYGETREYG 480
|
|
|
481  NFGSTRVYEAGHEVPYTOPTASLQFNRTIFGWDIAEGOKKIWPSTKNTGTATATHTQSS 540
|
|
|
481  NFGSTRVYEAGHEVPYTOPTASLQFNRTIFGWDIAEGOKKIWPSTKNTGTATATHTQSS 540
|
|
|
541  VLPPTATSMSSVGM 554
|
|
|
541  VLPPTATSMSSVGM 554
|
|
|

RESULT 2
US-08-943-714-9
; Sequence 9, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berka, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dambmann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; TITLE OF INVENTION: Encoding Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:

```



```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/309,341
; FILING DATE: 16-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus Niger
; US-08-309-341-2

Query Match
Best Local Similarity 21.1%; Score 630; DB 1; Length 557;
Matches 154; Conservative 81; Mismatches 174; Indels 76; Gaps 17;

QY 45 ANNVTIRKBPAGAGVCTTPGVKSYGYVDTSP--ESHTFFWFFFEARHNPETAPITLWLN 103
Db 132 AYDLRVKKTDPGSLG---DPGVKQYTGYYLDNENDKHLFYWFESRNDPNDPVVWLN 188

QY 104 GPGSDSLGLFEEELGPGCHVNSTFDDYINPHSWNEVSNLLFLSQPLGVGFYSYDVTGSI 163
Db 189 GPGCSSLTGLFELGPGSSINKKIQPVYNDYAWNSASVIFLDQPVNVGYSYNSA---- 244

QY 164 NPVTGVWENSSFAGVQGYPTIDATLIDTTLAAEAWEILQGLSLGSLPSLDSRVQSKDF 223
Db 245 -----VSDTVAAGKDVYALLTLFFKQFP-----EYAKQDF 274

QY 224 SLWTSYGGHYGPAFFNHFYEQNERIANGSVNGVQLNFNSLGIINGIIDEAIQAPYYPF 283
Db 275 HIAGESYAGHYIPVFASEILSHKKR-----NINLQSVLIGNGLTDGYTYEYRPM 325

QY 284 AVNNTYGIKAVNETVYNYMKFANQMPNGCQDLISTCKQTNRTALADYALCAEATNMCRDN 343
Db 326 ACGDGGYPAVLDESSCQMD--NALPR-CQSMIESCYSESA-----WVCVPASYCNA 377

QY 344 VEGPYAFAGRGVYDIRHPYDDP-----TPPSYNNKFLAKDSVMDAIGVIN--YTQSNND 397
Db 378 LLAPYQR-TGQNVYDVRGKCEDSSNLCYSAMGYVSDYLNKPEVIEAVGAENVGYDSCNFD 436

QY 398 VYAFQQTGDFWVP--NFIEDLEELALPVRVSLIY-GDADYICNWFQGOAVSLAANYSQ 454
Db 437 INRNFLPHGDMKPYHRLVPGILLE--QIPV---LIYAGDADFICNWLGNKAWTEALEWPG 491

QY 455 AAQFRSAGYTPLVK-----NGVEYGETREYGNFSFTRVYEAGHEVPYQPIASLQLENRT 509
Db 492 QAEVASAELEDLVVDNEHTGKKIQQVSHGNFTFMRLYGGHVMVPMQDPESSEFFENRW 551

QY 510 IFG-W 513
Db 552 LGGEW 556

```

RESULT 4

```

US-08-608-267-2
; Sequence 2, Application US/08608267
; Patent No. 5688663
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann

```

```

; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5688663o No. 5688663disk of No. 5688663th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,267
; FILING DATE: 28-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 557 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus Niger
; US-08-608-267-2

```

```

Query Match
Best Local Similarity 21.1%; Score 630; DB 1; Length 557;
Matches 154; Conservative 81; Mismatches 174; Indels 76; Gaps 17;

QY 45 ANNVTIRKBPAGAGVCTTPGVKSYGYVDTSP--ESHTFFWFFFEARHNPETAPITLWLN 103
Db 132 AYDLRVKKTDPGSLG---DPGVKQYTGYYLDNENDKHLFYWFESRNDPNDPVVWLN 188

QY 104 GPGSDSLGLFEEELGPGCHVNSTFDDYINPHSWNEVSNLLFLSQPLGVGFYSYDVTGSI 163
Db 189 GPGCSSLTGLFELGPGSSINKKIQPVYNDYAWNSASVIFLDQPVNVGYSYNSA---- 244

QY 164 NPVTGVWENSSFAGVQGYPTIDATLIDTTLAAEAWEILQGLSLGSLPSLDSRVQSKDF 223
Db 245 -----VSDTVAAGKDVYALLTLFFKQFP-----EYAKQDF 274

QY 224 SLWTSYGGHYGPAFFNHFYEQNERIANGSVNGVQLNFNSLGIINGIIDEAIQAPYYPF 283
Db 275 HIAGESYAGHYIPVFASEILSHKKR-----NINLQSVLIGNGLTDGYTYEYRPM 325

QY 284 AVNNTYGIKAVNETVYNYMKFANQMPNGCQDLISTCKQTNRTALADYALCAEATNMCRDN 343
Db 326 ACGDGGYPAVLDESSCQMD--NALPR-CQSMIESCYSESA-----WVCVPASYCNA 377

QY 344 VEGPYAFAGRGVYDIRHPYDDP-----TPPSYNNKFLAKDSVMDAIGVIN--YTQSNND 397
Db 378 LLAPYQR-TGQNVYDVRGKCEDSSNLCYSAMGYVSDYLNKPEVIEAVGAENVGYDSCNFD 436

QY 398 VYAFQQTGDFWVP--NFIEDLEELALPVRVSLIY-GDADYICNWFQGOAVSLAANYSQ 454
Db 437 INRNFLPHGDMKPYHRLVPGILLE--QIPV---LIYAGDADFICNWLGNKAWTEALEWPG 491

QY 455 AAQFRSAGYTPLVK-----NGVEYGETREYGNFSFTRVYEAGHEVPYQPIASLQLENRT 509

```

284 AVNTYGIKAVNETVYNYMKFANQMGCCODLLSTCKQTNRTADYALCAEATNMCRDN 343
 326 ACDDGGYPVAVLDESSQSDM--NALPR-COSMIESCYSSSA-----WVCVPASIIYCNNA 377
 344 VEGPYAFAGRGVYDIRHPYDDP-----TPPSYNNKELAKDSVMDAIGVNNIN-YTQSNND 397
 378 LLAPYQR-TGQNVYDVGKCEDSSNLCYSANGVYSDYLNKPEVIEAVGAENVGYDSCNFD 436
 398 VYFAFOOTGDFVWP--NFIEDLEILALPVRSIIY-GDADYICNWFGGQAVSLAANYSQ 454
 437 INRNFLPHGDMWKPEYHRLVPGLLF--QIPV---LIYAGDADFICNWLGNKAWTEALEWPG 491
 455 AAQFRSAGYPLKV-----NGVEYGETREYGNFSFTVYBAGHEVPYQPIASLQLFNRT 509
 492 QAEYASAELEDLVTDNEHTGKIGQVKSHGNFTFMRLYGGCHVMPMDQPESSLEFFNRW 551
 510 IFG-W 513
 552 LGGEW 556

RESULT 6
 US-08-608-224-2
 ; Sequence 2, Application US/08608224
 ; Patent No. 5705376
 ; GENERAL INFORMATION:
 ; APPLICANT: Yaver, Debbie Sue
 ; APPLICANT: Thompson, Sheryl Ann
 ; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: No. 57053760 No. 5705376disk of No. 5705376th America, Inc.
 ; STREET: 405 Lexington Avenue, Suite 6400
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/608,224
 ; FILING DATE: 28-FEB-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/309,341
 ; FILING DATE: 20-SEP-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney, Karen A.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 4247.000-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 867 0123
 ; TELEFAX: 212 867 0298
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 557 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; ORIGINAL SOURCE:
 ; ORGANISM: Aspergillus Niger
 ; US-08-608-224-2

Query Match 21.1%; Score 630; DB 1; Length 557;
 Best Local Similarity 31.8%; Pred. No. 7.2e-49;
 Matches 154; Conservative 81; Mismatches 174; Indels 76; Gaps 17;

45 ANNVITRYKEFGAGVCCTTPGVKSYGSDTSP-ESHTFFWFPEARHNPETAITLWLN 103
 132 AYDLRVKKTDPGSLGI---DFGVKQYTGILDDNENDKHLFWFSESDNDPVPVLMN 198
 104 GPGFSDSLIGLFLGELGCHVNSTFDDYINPHSMNEVSNLFLSQPLGVGFSYSDTVGSI 163
 189 GPGFCSLTLGLFELGFSINKKTPQVNDYAMNSASVIFLDQPVNVGYSYNSA---- 244
 164 NPVTGVVNSFAGVQGRYPTIDATLDTNLAARAAWEILQGLSLGSLDSRVQSKDF 223
 245 -----VSDTVAAGKVYALTLTFPKQFP-----EYAKQDF 274
 224 SLWTESYGGHYGPAFFNHFEQNERIANGSVNGVQLNFNSLGIINGIDEAIQAPYYPEF 283
 275 HIAGESYAGHYIPVFASEILLSKKR-----NINLQSVLIGNLTDGYQYVYRPM 325

492 QAEYASAELEDLVTDNEHTGKIGQVKSHGNFTFMRLYGGCHVMPMDQPESSLEFFNRW 551
 510 IFG-W 513
 552 LGGEW 556

RESULT 5
 US-08-608-452-2
 ; Sequence 2, Application US/08608452
 ; Patent No. 5693510
 ; GENERAL INFORMATION:
 ; APPLICANT: Yaver, Debbie Sue
 ; APPLICANT: Thompson, Sheryl Ann
 ; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: No. 56935100 No. 5693510disk of No. 5693510th America, Inc.
 ; STREET: 405 Lexington Avenue, Suite 6400
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/608,452
 ; FILING DATE: 28-FEB-1996
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/309,341
 ; FILING DATE: 20-SEP-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney, Karen A.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 4247.000-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 867 0123
 ; TELEFAX: 212 867 0298
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 557 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; ORIGINAL SOURCE:
 ; ORGANISM: Aspergillus Niger
 ; US-08-608-452-2

Query Match 21.1%; Score 630; DB 1; Length 557;
 Best Local Similarity 31.8%; Pred. No. 7.2e-49;
 Matches 154; Conservative 81; Mismatches 174; Indels 76; Gaps 17;

45 ANNVITRYKEFGAGVCCTTPGVKSYGSDTSP-ESHTFFWFPEARHNPETAITLWLN 103
 132 AYDLRVKKTDPGSLGI---DFGVKQYTGILDDNENDKHLFWFSESDNDPVPVLMN 198
 104 GPGFSDSLIGLFLGELGCHVNSTFDDYINPHSMNEVSNLFLSQPLGVGFSYSDTVGSI 163
 189 GPGFCSLTLGLFELGFSINKKTPQVNDYAMNSASVIFLDQPVNVGYSYNSA---- 244
 164 NPVTGVVNSFAGVQGRYPTIDATLDTNLAARAAWEILQGLSLGSLDSRVQSKDF 223
 245 -----VSDTVAAGKVYALTLTFPKQFP-----EYAKQDF 274
 224 SLWTESYGGHYGPAFFNHFEQNERIANGSVNGVQLNFNSLGIINGIDEAIQAPYYPEF 283
 275 HIAGESYAGHYIPVFASEILLSKKR-----NINLQSVLIGNLTDGYQYVYRPM 325

us-09-712-338-2.ra1

Mon Nov 24 13:41:08 2003

ADDRESSEE: No. 5688663c No. 5688663disk of No. 5688663th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/608,267
FILING DATE: 28-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/309,341
FILING DATE: 20-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4247.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 557 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Aspergillus Niger
US-08-608-267-4

Query Match 20.9%; Score 623; DB 1; Length 557;
Best Local Similarity 32.0%; Pred. No. 3.1e-48;
Matches 155; Conservative 79; Mismatches 175; Indels 76; Gaps 18;

QY	45	ANNVTIRKEPGAEGVCTTPGVKSYGVVDTP-ESHTFFWFFEARHNPETAPITLWLN	103
Db	132	AYDLVKKTKDPSLSGI---DPGVKQYTGVLDDNENDKHLFYWFFSRNDPNDPVVLWLN	188
QY	104	GGPGSDSLGLFEELGCPCHVNSTFDYINPHSWNEVSNLLFLSQPLGVGFSDTVDGSI	163
Db	189	GGPGCSLTGLFMELGPSSINKKIQPVNDYAWNSASVIFLDQPVNVGYSYNSA----	244
QY	164	NPVTGVVENSFAGVQGYPTIDATLIDTNTLAAEAWEILOGLFSLPSLSRSQKDF	223
Db	245	-----VSDTVAAGKDVYALLTLFFKQFP-----EYAKQDF	274
QY	224	SLWTSYGGHYGPAFFNHFYEQNERJANGSVQVNFNSLGIINGIIDEAIAQAPYPEF	283
Db	275	HIAGESYAGHYIPVFASEILSHKKR-----NINLSVLINGLTDGLTQYEYRPM	325
QY	284	AVNNTYGIKAVNETVYNYMKFANQMPNGCQDLSTCKQNTRTALADYALCAEATNCRDN	343
Db	326	ACDGD-GYPVAVLDE-GSCQAMDNALPR-CQSMIESCVSSEA-----WVCVPASIYCNNA	377
QY	344	VEGPIYAFAGRGVYDIRHPYDDP-----TPPSYNNKFLAKDSYMDAIGVNNIN-YTQSNND	397
Db	378	LLAPYQR-TQONVYDVRGKCEDSSNLCYSAMGVSDYLNTKTEVIEAVGAENVGYDSCNFD	436
QY	398	VYVAFQGTDFVMP--NFIEDLEILALPVRVSLIY--GDADYICNWFQGGQAVSLAANSYQ	454
Db	437	INRNFLPHGDMKPYHRLVPGLE--QIPV--LIVAGDADFICNWLGNKAWTEALEWPG	491
QY	455	AAQFRSAGYTPLVK-----NGVEYGETREYGNFSFTRVYAGHEVPYQYIASIQLPNRT	509
Db	492	QAEYASAKLEDLVVENEHKKIGQVKSHGNTFTMELYGCGHVPMDQPESSLEFFNRW	551
QY	510	IFG-W 513	
Db	552	LGGEW 556	

RESULT 9
US-08-608-267-4
Sequence 4, Application US/08608267
Patent No. 5688663
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie Sue
APPLICANT: Thompson, Sheryl Ann
TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

QY	45	ANNVTIRKEPGAEGVCTTPGVKSYGVVDTP-ESHTFFWFFEARHNPETAPITLWLN	103
Db	132	AYDLVKKTKDPSLSGI---DPGVKQYTGVLDDNENDKHLFYWFFSRNDPNDPVVLWLN	188
QY	104	GGPGSDSLGLFEELGCPCHVNSTFDYINPHSWNEVSNLLFLSQPLGVGFSDTVDGSI	163
Db	189	GGPGCSLTGLFMELGPSSINKKIQPVNDYAWNSASVIFLDQPVNVGYSYNSA----	244
QY	164	NPVTGVVENSFAGVQGYPTIDATLIDTNTLAAEAWEILOGLFSLPSLSRSQKDF	223
Db	245	-----VSDTVAAGKDVYALLTLFFKQFP-----EYAKQDF	274
QY	224	SLWTSYGGHYGPAFFNHFYEQNERJANGSVQVNFNSLGIINGIIDEAIAQAPYPEF	283
Db	275	HIAGESYAGHYIPVFASEILSHKKR-----NINLSVLINGLTDGLTQYEYRPM	325
QY	284	AVNNTYGIKAVNETVYNYMKFANQMPNGCQDLSTCKQNTRTALADYALCAEATNCRDN	343
Db	326	ACDGD-GYPVAVLDE-GSCQAMDNALPR-CQSMIESCVSSEA-----WVCVPASIYCNNA	377
QY	344	VEGPIYAFAGRGVYDIRHPYDDP-----TPPSYNNKFLAKDSYMDAIGVNNIN-YTQSNND	397
Db	378	LLAPYQR-TQONVYDVRGKCEDSSNLCYSAMGVSDYLNTKTEVIEAVGAENVGYDSCNFD	436
QY	398	VYVAFQGTDFVMP--NFIEDLEILALPVRVSLIY--GDADYICNWFQGGQAVSLAANSYQ	454
Db	437	INRNFLPHGDMKPYHRLVPGLE--QIPV--LIVAGDADFICNWLGNKAWTEALEWPG	491
QY	455	AAQFRSAGYTPLVK-----NGVEYGETREYGNFSFTRVYAGHEVPYQYIASIQLPNRT	509
Db	492	QAEYASAKLEDLVVENEHKKIGQVKSHGNTFTMELYGCGHVPMDQPESSLEFFNRW	551
QY	510	IFG-W 513	

Db 552 LGGEW 556

RESULT 10

US-08-608-452-4

Sequence 4, Application US/08608452

Patent No. 5693510

GENERAL INFORMATION:

APPLICANT: Yaver, Debbie Sue

APPLICANT: Thompson, Sheryl Ann

TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 56935100 No. 5693510disk of No. 5693510th America, Inc.

STREET: 405 Lexington Avenue, Suite 6400

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/608,452

FILING DATE: 28-FEB-1996

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/309,341

FILING DATE: 20-SEP-1994

NAME: Lowney, Karen A.

REGISTRATION NUMBER: 31,274

REFERENCE/DOCKET NUMBER: 4247.000-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 867 0123

TELEFAX: 212 867 0298

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 557 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Aspergillus Niger

US-08-608-452-4

Query Match

Best Local Similarity 20.9%; Score 623; DB 1; Length 557;

Matches 155; Conservative 79; Mismatches 175; Indels 76; Gaps 18;

QY 45 ANNVTIRKPEAGVCCTTPGVKSYSGYVDTP-ESHTFFWFPEARHNPTAPITLWLN 103

Db 132 AYDLRVKKTDPSSLGI---DPGVKQYTGYYLDNDNDKHLFYWFESRNDPNDPVVLN 188

QY 104 GPGSDSLIGLFEELGPGCHVNSTPDDYINPHSNWNEVSNLLFLSQPLGVGFYSYSDTVGSI 163

Db 189 GPGCSSLTGLFMELGPGSSINKKIOPVYNDVWANSNASVIFLDQPVNVGYSYNSA---- 244

QY 164 NPVTGVNENSSFAGVQGRYPTIDATLIIDTNLAEEAWEILQGLPSLDNRVQSKDF 223

Db 245 -----VSDTVAGKDVYALLTLFPKQP-----EVAKQDF 274

QY 224 SLWTESYGGHGPFAFFNPFQNERIANGSVNGVOLFNSLGIINGIIDEAIQAPYYPEF 283

Db 275 HIAGESYAGHYIPVFASEILSHKKR-----NINLQSVLIGNGLDGLTQVEYYRPM 325

QY 284 AVNNTYGIKAVNETVYNNKFAQNPNGCODLISTCKQKTNTALADYALCAEATMCRDN 343

Db 326 ACGDG-GYPAVLDE-GSCQAMDNALPR-CQSMIESCSYSSA-----WVCVPASTYCNNA 377

QY 344 VEGPYAFAGRGVYDIRHPYDDP-----TPPSYNNKELAKDSVMDAIGVNNIN-YTQSNND 397

Db 378 LLAPYQR-TGQNVYDVRGKCEDSSNLCSAMGYVSDYLNKTEVIEAGVNGYDSCNFD 436

QY 398 VYVAFQOTGDPVMP--NFIEDLEILALPVRVSLIY-GDADYICNWFPGGQAVSLAANSQ 454

Db 437 INRNFLPHGDMKPYHRLVPLLE--QIPV---LIYAGDADFICNMLGNKAWTALEWPG 491

QY 455 AAQFRSAGYTPLKV-----NGVYGETREYGNFSFTRVYEAGHEVPYQYPIASLOLENRT 509

Db 492 QAEVASAKLEDVAVNEHKGKIGQVKSHGNFTFMRLYCGGHWVPMQDPESSLEFFNRW 551

QY 510 IFG-W 513

Db 552 LGGEW 556

RESULT 11

US-08-608-224-4

Sequence 4, Application US/08608224

Patent No. 5705376

GENERAL INFORMATION:

APPLICANT: Yaver, Debbie Sue

APPLICANT: Thompson, Sheryl Ann

TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 57053760 No. 5705376disk of No. 5705376th America, Inc.

STREET: 405 Lexington Avenue, Suite 6400

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/608,224

FILING DATE: 28-FEB-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/309,341

FILING DATE: 20-SEP-1994

NAME: Lowney, Karen A.

REGISTRATION NUMBER: 31,274

REFERENCE/DOCKET NUMBER: 4247.000-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 867 0123

TELEFAX: 212 867 0298

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 557 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Aspergillus Niger

US-08-608-224-4

Query Match

Best Local Similarity 20.9%; Score 623; DB 1; Length 557;

Matches 155; Conservative 79; Mismatches 175; Indels 76; Gaps 18;

QY 45 ANNVTIRKPEAGVCCTTPGVKSYSGYVDTP-ESHTFFWFPEARHNPTAPITLWLN 103

Db 132 AYDLRVKKTDPSSLGI---DPGVKQYTGYYLDNDNDKHLFYWFESRNDPNDPVVLN 188

QY 104 GPGSDSLIGLFEELGPGCHVNSTPDDYINPHSNWNEVSNLLFLSQPLGVGFYSYSDTVGSI 163

Db 189 GPGCSSLTGLFMELGPGSSINKKIOPVYNDVWANSNASVIFLDQPVNVGYSYNSA---- 244

QY 164 NPVTGVNENSSFAGVQGRYPTIDATLIIDTNLAEEAWEILQGLPSLDNRVQSKDF 223

Db 245 -----VSDTVAGKDVYALLTLFPKQP-----EVAKQDF 274

QY 224 SLWTESYGGHGPFAFFNPFQNERIANGSVNGVOLFNSLGIINGIIDEAIQAPYYPEF 283

Db 275 HIAGESYAGHYIPVFASEILSHKKR-----NINLQSVLIGNGLDGLTQVEYYRPM 325

QY 284 AVNNTYGIKAVNETVYNNKFAQNPNGCODLISTCKQKTNTALADYALCAEATMCRDN 343

Db 326 ACGDG-GYPAVLDE-GSCQAMDNALPR-CQSMIESCSYSSA-----WVCVPASTYCNNA 377

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Aspergillus Niger
US-08-967-149-4

Query Match
Best Local Similarity 32.0%; Pred. No. 3.1e-48;
Matches 155; Conservative 79; Mismatches 175; Indels 76; Gaps 18;

Qy 45 ANVTIYKPGAGVCETTPGVKSYGYVDTSP-ESHTEFWFPEARHNPETAPITLWLN 103
Db 132 AYDLRVKTPFPSSLGI---DPGVKQYTGYYLDNDNDKHLFWWFESRNDPDPVWLWLN 188
Qy 104 GPGCSLSLIGLFEELGCHVNSTPDDYINPHSMNEVSNLLFSLQPLGVGFYSYDTPVDSI 163
Db 189 GPGCSLSLTGLFMELGFSSINNKIQPVYNDVAMNSNASVIFLDQPVNVGYSYNSA---- 244
Qy 164 NPVTGVVNSSFAGVQGRYPTIDATLDTTNLAEEAAWEILQGLSLGLSDSRVQSKDF 223
Db 245 -----VSDTVAAGKDVYALLTLFFKQFP-----EYAKQDF 274
Qy 224 SLWTESYGGHYGPAFFNFHFEQNERIANGSVNGVQLNENSLGIINGIIDEAIQAPYPEF 283
Db 275 HIAGESYAGHYIPVFASLILSHKXR-----NINLQSVLIGNGLTDGLTQYEYRPM 325
Qy 284 AVNTYGIKAVNETVYNYMKFANQMPNGCQDLISTCKQTNRTALADYALCAEATNWCND 343
Db 326 ACGDG-GYPAVLDE-GSCQAMDNALPR-CQSMIESCYSSA-----WVCVPASTYCNNA 377
Qy 344 VEGPIYAFAGRGVDIRHPYDDP-----TPPSYNNKFLAKOSVMDAIGVNNIN-VTQSNND 397
Db 378 LLAPYQR-TGQNVYDVRGKCEDSSNLCSAMGYVSDYLNTKTEVIEAAGVNGYDSCNFD 436
Qy 398 VYVAFQQTGDFWMP-NFIEDLEETALPVRVSLIY-GDADYICNWFQGGQAVSLAANYSQ 454
Db 437 INRNFLFGDNWPKYHRLVPGLE--QIPV---LIYAGDADFICNWLGNKAWTEALEWPG 491
Qy 455 AAQFRSAGYTPLKV-----NGVEYGETREYNFSTRVYEAGHEVPYIOPIASLQIFNRT 509
Db 492 QAEYASAKLEDLVVVENEHKKGKIGQVKSHGNFTFMRLYGGGHVMPMDQPESSLEFFNRW 551
Qy 510 IFG-W 513
Db 552 LGGEW 556

RESULT 13
US-09-640-305-4
Sequence 4, Application US/09640305
Patent No. RE37447
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Yeh, Patrice
TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
PREPARATION AND USE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd. 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19002
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

189 GPGCSLSLTGLFMELGFSSINNKIQPVYNDVAMNSNASVIFLDQPVNVGYSYNSA---- 244
Qy 164 NPVTGVVNSSFAGVQGRYPTIDATLDTTNLAEEAAWEILQGLSLGLSDSRVQSKDF 223
Db 245 -----VSDTVAAGKDVYALLTLFFKQFP-----EYAKQDF 274
Qy 224 SLWTESYGGHYGPAFFNFHFEQNERIANGSVNGVQLNENSLGIINGIIDEAIQAPYPEF 283
Db 275 HIAGESYAGHYIPVFASLILSHKXR-----NINLQSVLIGNGLTDGLTQYEYRPM 325
Qy 284 AVNTYGIKAVNETVYNYMKFANQMPNGCQDLISTCKQTNRTALADYALCAEATNWCND 343
Db 326 ACGDG-GYPAVLDE-GSCQAMDNALPR-CQSMIESCYSSA-----WVCVPASTYCNNA 377
Qy 344 VEGPIYAFAGRGVDIRHPYDDP-----TPPSYNNKFLAKOSVMDAIGVNNIN-VTQSNND 397
Db 378 LLAPYQR-TGQNVYDVRGKCEDSSNLCSAMGYVSDYLNTKTEVIEAAGVNGYDSCNFD 436
Qy 398 VYVAFQQTGDFWMP-NFIEDLEETALPVRVSLIY-GDADYICNWFQGGQAVSLAANYSQ 454
Db 437 INRNFLFGDNWPKYHRLVPGLE--QIPV---LIYAGDADFICNWLGNKAWTEALEWPG 491
Qy 455 AAQFRSAGYTPLKV-----NGVEYGETREYNFSTRVYEAGHEVPYIOPIASLQIFNRT 509
Db 492 QAEYASAKLEDLVVVENEHKKGKIGQVKSHGNFTFMRLYGGGHVMPMDQPESSLEFFNRW 551
Qy 510 IFG-W 513
Db 552 LGGEW 556

RESULT 12
US-08-967-149-4
Sequence 4, Application US/08967149
Patent No. 5939305
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie Sue
APPLICANT: Thompson, Sheryl Ann
TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF
TITLE OF INVENTION: ASPERGILLUS NIGER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5939305 No. 5939305disk of No. 5939305th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,149
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,452
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 08/309,341
FILING DATE: 20-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowrey, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4247.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 557 amino acids

```

; APPLICATION NUMBER: US/09/640,305
; FILING DATE: 16-Aug-2000
; PRIORITY DATA:
; APPLICATION NUMBER: US/08/360,673
; FILING DATE: 06-FEB-1995
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-640-305-4

Query Match
Best Local Similarity 19.1%; Score 568.5; DB 1; Length 491;
Matches 151; Conservative 66; Mismatches 184; Indels 85; Gaps 14;

QY 45 ANNVTIRYKEPGAEGVCEITPGVKSYGVYDTSPESHTEFWFEARHNPTAPITLWNG 104
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 AYSLRIKPLDPKSLGV-----DTVKQWSGYLDYQDSKHFFYWFESRNDPNDPVILWNG 117
QY 105 GPGSDSLIGLFEELGPGCHVNSTFDDYINPHSWNEVSNLLFSLQPLGVGFSYSDTVDSIN 164
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
118 GPGCSFVGLFPELGFSSIGADLKPIYNPYSWNSNASVIFLDQPVGVGFSYGD----- 170
QY 165 PVTGVVENSFAGVQGRYPTDITDITNLAEEAAWEILQGLSLGLPSLDSRVQSKDFS 224
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
171 -----SKVSTTDDAAKDVYIFLDLFFERFPHL-----RNNDFH 203
QY 225 LMTESYGHYGPAPFNHFEQNERIANGSVNGVQLNFNSLGIINGIIDEAIQAPYYPEFA 284
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
204 IGSYAGHYLPKIAH-----EIAVVAEDSSFNLSVLINGFTDPLTQYQYEPMA 256
QY 285 VNNYGIKAVNETVYNYMKFANMNGCOD-----LISTCKQTNRTALADYALCA 334
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
257 CGEG-GYPAVLE-----PEDCLDMNRNLPCLSLVDRCKYKSH-----SVFSCV 298
QY 335 EATNCRDNVEGYPYAFAGRGVYDIR---HPYDDP-----TPPSYNNKFLAKDSVMDAIGV 387
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
299 LADRYCEQQITG-VYEKSGRNPYDIRSKCEAEDDSGACYQEEIYISDYLNOBEVQALGT 357
QY 388 NINYTQS--NNDVYVAFQQTGDFWPNFIEDLEIILALPVRVSLYGDADYICNWFQGGQAV 446
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
358 DVSSFQGCSSDVIGIGFAGTGPSP-FHQYVAELLDDQDINVLVYAGDKDYICNWLGNLAW 416
QY 447 SLAANSQAQPSRAGYTPLVKNGVE--YGETREYNGFSFTVYEAGHEVPYQPIASLO 504
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
417 TEKLEWRYNEEYKQVLRWKSEETDETIGETKSYGPTLYLRIYDAGHVMVHDQENSILQ 476
QY 505 LPNRTI 510
Db : : : : :
477 MVNSWI 482

```

RESULT 14

```

US-08-360-673-4
; Sequence 4, Application US/08360673
; Patent No. 5679544
; GENERAL INFORMATION:
; APPLICANT: Fleet, Reinhard
; APPLICANT: Fournier, Alain

```

```

; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,673
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-360-673-4

```

```

Query Match
Best Local Similarity 19.1%; Score 568.5; DB 1; Length 491;
Matches 151; Conservative 66; Mismatches 184; Indels 85; Gaps 14;

QY 45 ANNVTIRYKEPGAEGVCEITPGVKSYGVYDTSPESHTEFWFEARHNPTAPITLWNG 104
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 AYSLRIKPLDPKSLGV-----DTVKQWSGYLDYQDSKHFFYWFESRNDPNDPVILWNG 117
QY 105 GPGSDSLIGLFEELGPGCHVNSTFDDYINPHSWNEVSNLLFSLQPLGVGFSYSDTVDSIN 164
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
118 GPGCSFVGLFPELGFSSIGADLKPIYNPYSWNSNASVIFLDQPVGVGFSYGD----- 170
QY 165 PVTGVVENSFAGVQGRYPTDITDITNLAEEAAWEILQGLSLGLPSLDSRVQSKDFS 224
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
171 -----SKVSTTDDAAKDVYIFLDLFFERFPHL-----RNNDFH 203
QY 225 LMTESYGHYGPAPFNHFEQNERIANGSVNGVQLNFNSLGIINGIIDEAIQAPYYPEFA 284
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
204 IGSYAGHYLPKIAH-----EIAVVAEDSSFNLSVLINGFTDPLTQYQYEPMA 256
QY 285 VNNYGIKAVNETVYNYMKFANMNGCOD-----LISTCKQTNRTALADYALCA 334
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
257 CGEG-GYPAVLE-----PEDCLDMNRNLPCLSLVDRCKYKSH-----SVFSCV 298
QY 335 EATNCRDNVEGYPYAFAGRGVYDIR---HPYDDP-----TPPSYNNKFLAKDSVMDAIGV 387
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
299 LADRYCEQQITG-VYEKSGRNPYDIRSKCEAEDDSGACYQEEIYISDYLNOBEVQALGT 357
QY 388 NINYTQS--NNDVYVAFQQTGDFWPNFIEDLEIILALPVRVSLYGDADYICNWFQGGQAV 446
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
358 DVSSFQGCSSDVIGIGFAGTGPSP-FHQYVAELLDDQDINVLVYAGDKDYICNWLGNLAW 416

```

QY 447 SLAANYQAQFRSAGYTPKLVNGVE--YGETREYGNFSFTRVYEAHEVYVYQPIASLQ 504

Db 417 TEKLEWRYNEEYKQVLRTWKSEETDETIGTKSYGLTYLRIYDAGHMVPHDQPENSLQ 476

QY 505 LENRTI 510

Db 477 MVNSWI 482

RESULT 15

US-08-899-324-33

Sequence 33, Application US/08899324

Patent No. 5945329

GENERAL INFORMATION:

APPLICANT: Breddam, Klaus

APPLICANT: Keilland-Brandt, Morten

APPLICANT: Mortensen, Uffe

APPLICANT: Olesen, Kjeld

APPLICANT: Stennicke, Henning

APPLICANT: Wagner, Fred

TITLE OF INVENTION: CUSTOMIZED PROTEASES

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt

STREET: 3100 No. 5945329west Center, 90 S. 7th Street

CITY: Minneapolis

STATE: MN

COUNTRY: U.S.A.

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/899,324

FILING DATE: 23-JUL-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/329,892

FILING DATE: 27-OCT-1994

APPLICATION NUMBER: 08/144,704

FILING DATE: 28-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Kettleberger, Denise M

REGISTRATION NUMBER: 33,924

REFERENCE/DOCKET NUMBER: 8648.44USC1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612/332-5300

TELEFAX: 612/332-9081

TELEX:

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 532 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

US-08-899-324-33

Query Match 17.6%; Score 523.5; DB 2; Length 532;

Best Local Similarity 29.1%; Pred. No. 3.6e-39;

Matches 147; Conservative 74; Mismatches 200; Indels 85; Gaps 19;

QY 33 PKNFTGVKT-----LTTANNVTIRY-----KEFGAGVCEITTPGVKSYSGYVDTSPF- 79

Db 82 PKFPEAIKTKDWDFFVVKNDAIENYQLRVNKKDKPKILGI---DPNVTQYTGLDVDEDE 138

QY 80 SHTFWFPEARHNPETAPITLNLNGPGSDSLIGLFEELGCPCHVNSTFDDYINPHSWNEV 139

Db 139 KHFFFTFESRNDPAKDPVILNLNGGFGCSSLTGLFFELGSSIGPDLKPTIGNFYNSKN 198

QY 140 SNLLFLSQPLGVSFYSYSDTVDSINPVTGVVENSFPAGVQGRYPTIDATLIDTTNLAAEA 199

Db 199 ATVIFLDQPVNVGFSYS-----GSSG-----VSNTVAAGKD 229

QY 200 AWEILOGLSGLPSLDSRVQSKDFSLWMTESYGGHYGPAFFNHFYEQNERIANGSVNGVQL 259

Db 230 VYNFELFLFDQFPPEYVKNQG--DPIHAGESYAGHYIPVFAGEILSHKDR----- 276

QY 260 NFNSLGI--NGIIDEAIAQAPYEPFAVNNITYGKAVNETVYNYMKFANQMPNGCQDLIS 317

Db 277 NFNLTSVLINGLIDPLTQYNYEPMACGEG-GEPSV-----LPSECSAMEDSLE 326

QY 318 TCKQNTALADYAL--CAEATNMCNDRNVEGPFYAFAGRGYVDIRHPYDD-----PTPFS 370

Db 327 RCLGLIESSYDSQSVWSCVPATIIYCNNAQLAPYQR-TGRNVDYDIRKDCGEGNLCYPTIQD 385

QY 371 YNKFELAKDSYMDAIGVNNINYTQS--NNDVYVAFQQTGDFVWPNFTEDLEEILALPVRVSL 429

Db 386 -IDYLNQDYVYKBAVGAEDVHYESCNFDINRNFLFAGDWMKP-YHTAVTDLNLQDLPIIV 443

QY 430 IYGDADYICNWFEGGQAVSLAANYSOAQFRS---AGYTPLKVNQVGEYGETREYGNFSFTR 486

Db 444 YAGDKDFICNWLGNKAWTDVLPWKYDEBFASQKVRNWTASITDEVA-GEVKSYPKHTVLR 502

QY 487 VYEAGHEVYFYQPIASLQLENRTIFG 512

Db 503 VENGGHMVFPDVPENALSMVNEWIHG 528

Search completed: November 21, 2003, 18:02:07

Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 17:51:55 ; Search time 46 Seconds
(without alignments)
1915.071 Million cell updates/sec

Title: US-09-712-338-2
Perfect score: 2979
Sequence: 1 MRGVEFLSVLPVLAASWALP.....HTQSSVPLPTATSMSSVQMA 555

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2975	99.9	554	19	AAW56099
2	2376	79.8	551	23	ABR38864
3	980.5	32.9	523	23	ABR38865
4	941	31.6	554	23	ABR38859
5	925	31.1	526	23	ABR38817
6	636	21.3	508	23	ABR38819
7	630	21.1	536	23	ABR38819
8	630	21.1	556	17	AA96737
9	620	20.8	557	17	AA96738
1	2975	99.9	554	19	AAW56099
2	2376	79.8	551	23	ABR38864
3	980.5	32.9	523	23	ABR38865
4	941	31.6	554	23	ABR38859
5	925	31.1	526	23	ABR38817
6	636	21.3	508	23	ABR38819
7	630	21.1	536	23	ABR38819
8	630	21.1	556	17	AA96737
9	620	20.8	557	17	AA96738

10	568.5	19.1	491	15	AA48059	Sequence of protea
11	541.5	18.2	550	23	ABG93281	C. albicans BAX-as
12	419	14.1	481	23	ABR38845	A. niger carboxype
13	383	12.9	482	21	AAG30065	Arabidopsis thalia
14	383	12.9	502	21	AAG30064	Arabidopsis thalia
15	365	12.3	455	23	ABR38843	A. niger carboxype
16	361.5	12.1	479	21	AAG31905	Arabidopsis thalia
17	361.5	12.1	486	21	AAG31904	Arabidopsis thalia
18	361	12.1	476	19	AAW72966	Human serine carbo
19	361	12.1	476	22	ABR88381	Human membrane or
20	359	12.1	476	20	AA133372	Amino acid sequenc
21	359	12.1	476	21	ABR1407	Human TANGO 176.
22	359	12.1	476	22	AAU29228	Human PRO polypept
23	359	12.1	476	22	ABR88587	Human hydrophobic
24	359	12.1	476	22	ABR80240	Human PRO223 prote
25	359	12.1	476	24	ABU69650	Novel human secret
26	359	12.1	476	24	ABU71316	Human PRO polypept
27	359	12.1	476	24	ABU71473	Human secreted/tra
28	359	12.1	476	24	ABU71919	Human secreted/tra
29	359	12.1	476	24	ABU65773	Novel human secret
30	359	12.1	476	24	ABU66106	Human secreted/pro
31	359	12.1	476	24	ABU67373	Human secreted/tra
32	359	12.1	476	24	ABU67610	Human secreted/pro
33	359	12.1	476	24	ABU64527	Human secreted/tra
34	359	12.1	476	24	ABU65468	Human secreted/tra
35	359	12.1	476	24	ABU58604	Human PRO polypept
36	359	12.1	476	24	ABU56140	Human PRO polypept
37	359	12.1	476	24	ABU57135	Human secreted/tra
38	359	12.1	476	24	ABU54375	Human PRO polypept
39	359	12.1	476	24	ABU10714	Human secreted/tra
40	359	12.1	477	22	AAU25810	Human secreted/tra
41	357.5	12.0	479	21	AAU47178	Human protein sequ
42	357.5	12.0	486	21	AAU47177	Arabidopsis thalia
43	357	12.0	476	20	AAU28570	Secreted peptide c
44	357	12.0	476	22	AAU39043	Human secreted pro
45	357	12.0	476	23	ABR55752	Human polypeptide

ALIGNMENTS

RESULT 1
AAW56099
ID AAW56099 standard; Protein: 554 AA.
XX
AC AAW56099;
XX
DT 27-AUG-1998 (first entry)
XX
DE A. oryzae ATCC20386 carboxypeptidase I protein.
XX
KW Carboxypeptidase I; flavour improving agent; hydrolysat; proteinaceous;
XX food industry.
XX
OS Aspergillus oryzae.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT Protein /label= signal
FT /label= 19..554
FT /label= carboxypeptidase I
XX
PN WO9814599-A1.
XX
PD 09-APR-1998.
XX
PF 03-OCT-1997; 97WO-US17977.
XX
PR 27-NOV-1996; 96US-0757534.
XX 04-OCT-1996; 96US-0726880.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO-NORDISK AS.

DR P-PSDB; ABZ78231, ABZ78288.

XX Novel isolated protease polypeptide useful in laboratory, clinical, pharmaceutical, chemical, diagnostic, personal care and industrial applications

XX Claim 13; Page 379-382; 394pp; English.

XX The invention relates to a novel isolated protease polypeptide. A polypeptide or polynucleotide of the invention is useful for diagnosing a fungal infection such as aspergillosis, or as a query sequence to perform a search against public databases. A polypeptide of the invention is useful in a selected number of industrial or pharmaceutical processes, in laboratory or clinical processes, in food industry (baking, brewing, cheese manufacturing, meat tenderising), in tanning industry and in the manufacture of biological detergents. A polypeptide may also be useful for improving protein solubility, extraction yields, viscosity or taste, texture, nutritional value, minimising of antigenicity or anti-nutritional factors, colour or functionality as well as processing aspects like filterability of the proteinaceous raw material. The sequences shown in ABR38865-ABR38869 represent the A. niger proteases of the invention.

XX SQ Sequence 551 AA;

Query Match 79.8%; Score 2376; DB 23; Length 551;
Best Local Similarity 78.1%; Pred. No. 9.6e-200;
Matches 428; Conservative 54; Mismatches 66; Indels 0; Gaps 0;

QY 1 MRGVEFLSLPLVAASWALPGSTPASVGRQLKPNFTGVTLTANNVTIRYKPGAGV 60
Db 1 MRGSRVLLPLAALSCAMPENESWSTIRQLKASTGVKSIKTPNNVTIRYKPGTEGI 60

QY 61 CETTPGVKSYGVYDTPSPESHTFFWFPEARHNPETAPITLWNGGPGSDSLGLFELGP 120
Db 61 CETTPGVKSYGVYDTPSPESHTFFWFPEARHNPETAPITLWNGGPGSDSLGLFELGP 120

QY 121 CHVNSTFDYINPHSNWENSLFLSQPLGVGFSYSDTVDGSIINPVTVGVNSSFAGVQG 180
Db 121 CHITPEYSIINQYSWNEVTNLLFLSQPLGVGFSYSETEAGSLNPTGAVENASFGVQG 180

QY 181 RYPTIDATLIDTNLAEEAWEILQGLSLPSLDSRVOSKPSLWTSYGGHYGPAFFN 240
Db 181 RYPTIDATLIDTNLAEEAWEILQGLSLPSLDSRVOSKPSLWTSYGGHYGPAFFN 240

QY 241 HFEQNERIANGSVQVLFNFSLSLGIINGIIDEAICAPYEPFAVNTYGIKAVNETVYN 300
Db 241 HFEQNSKIASEGVNGVQLNFSLSLGIINGIIDEAICAPYEPFAVNTYGIKAVNETVYN 300

QY 301 YMKFANOMPGCODLISTCKQNTALADYALCAEATNCRDNVEGFFYAFAGRGVYDIR 360
Db 301 YMKFANTWENGQDQVASCNLTNRSLSDYAICTEAAANCRDNVEGFFYAFAGRGVYDIR 360

QY 361 HPYDDPTPSYNNKFLAKDSVMDALGVNINVTQSNVDVYAFQQTGDFWPNFIEDLEI 420
Db 361 HPYNPTPSYFVDFLKKDSVMDALGVNINVTQSNVDVYAFQQTGDFWPNFIEDLEI 420

QY 421 LALPVRVSLIYGDADYICNWFQGVQVSLAANYSQAAQPSAGVTPKLVNGVYGETREYG 480
Db 421 LQLPVRVSLIYGDADYICNWFQGVQVSLAANYSQAAQPSAGVTPKLVNGVYGETREYG 480

QY 481 NFSFTRVTEAGHEVYQPIASQLFNRTIFGWDIAEGOKIWPYSYKNGTATATHTOSS 540
Db 481 NFSFTRVYQAGHEVYQPIAALQLFNRTLFGWDIAAGTQIWPYSYKNGTATHTESF 540

QY 541 VPLPTATS 548
Db 541 VPLSTASS 548

RESULT 3
ABR38865
ID ABR38865 standard; Protein; 623 AA.

XX AC ABR38865;
XX DT 24-APR-2003 (first entry)
XX DE A. niger serine carboxypeptidase polypeptide #4.
XX KW Protease; fungal infection; aspergillosis; food; tanning; detergent;
XX KW protein solubility; viscosity; taste; texture; nutritional value;
XX OS Aspergillus niger.
XX PN WO200268623-A2.
XX PD 06-SEP-2002.
XX PF 22-FEB-2002; 2002WO-EP01984.
XX PR 23-FEB-2001; 2001EP-0200657.
XX PR 23-FEB-2001; 2001EP-0200658.
XX PR 23-FEB-2001; 2001EP-0200660.
XX PR 26-FEB-2001; 2001EP-0200706.
XX PR 26-FEB-2001; 2001EP-0200707.
XX PR 26-FEB-2001; 2001EP-0200708.
XX PR 26-FEB-2001; 2001EP-0200719.
XX PR 28-MAR-2001; 2001EP-0000075.
XX PR 28-MAR-2001; 2001EP-0000078.
XX PR 28-MAR-2001; 2001EP-0000080.
XX PR 28-MAR-2001; 2001EP-0000087.
XX PR 28-MAR-2001; 2001EP-0000088.
XX PR 21-MAY-2001; 2001EP-0000156.
XX PR 21-MAY-2001; 2001EP-0000160.
XX PR 21-MAY-2001; 2001EP-0000162.
XX PR 21-MAY-2001; 2001EP-0000165.
XX PR 21-MAY-2001; 2001EP-0000166.
XX PR 21-MAY-2001; 2001EP-0000168.
XX PR 21-JUN-2001; 2001EP-0000240.
XX PR 21-JUN-2001; 2001EP-0000242.
XX PR 21-JUN-2001; 2001EP-0000244.
XX PR 21-JUN-2001; 2001EP-0000246.
XX PR 12-JUL-2001; 2001EP-0000280.
XX PR 12-JUL-2001; 2001EP-0000285.
XX PR 30-JUL-2001; 2001EP-0000323.
XX PR 02-AUG-2001; 2001EP-0000327.
XX PR 02-AUG-2001; 2001EP-0000341.
XX PR 02-AUG-2001; 2001EP-0000342.
XX PR 02-AUG-2001; 2001EP-0000343.
XX PR 02-AUG-2001; 2001EP-0000344.
XX PR 09-AUG-2001; 2001EP-0000357.
XX PR 16-AUG-2001; 2001EP-0000374.
XX PR 16-AUG-2001; 2001EP-0000377.
XX PR 20-SEP-2001; 2001EP-0000478.
XX PR 20-SEP-2001; 2001EP-0000483.
XX PR 22-OCT-2001; 2001EP-0000552.
XX PR 22-OCT-2001; 2001EP-0000553.
XX PR 22-OCT-2001; 2001EP-0000554.
XX PR 22-OCT-2001; 2001EP-0000556.
XX PR 22-OCT-2001; 2001EP-0000557.
XX PR 22-OCT-2001; 2001EP-0000558.
XX PR 15-NOV-2001; 2001EP-0004464.
XX PR 21-DEC-2001; 2001EP-0005117.
XX (STAM) DSM NV.
XX Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
XX Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
XX Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
XX Stiebler J, Albang R;
XX WPI; 2002-723203/78.
XX P-PSDB; ABZ78232, ABZ78289.

Mon Nov 24 13:41:08 2003

us-09-712-338-2.rag

XX Novel isolated protease polypeptide useful in laboratory, clinical, pharmaceutical, chemical, diagnostic, personal care and industrial applications

PS Claim 13; Page 382-385; 394pp; English.

XX The invention relates to a novel isolated protease polypeptide. A polypeptide or polynucleotide of the invention is useful for diagnosing a fungal infection such as aspergillosis, or as a query sequence to perform a search against public databases. A polypeptide of the invention is useful in a selected number of industrial or pharmaceutical processes, in laboratory or clinical processes, in food industry (baking, brewing, cheese manufacturing, meat tenderising), in tanning industry and in the manufacture of biological detergents. A polypeptide may also be useful for improving protein solubility, extraction yields, viscosity or taste, texture, nutritional value, minimising of antigenicity or anti-nutritional factors, colour or functionality as well as processing aspects like filterability of the proteinaceous raw material. The sequences shown in ABR3813-ABR3869 represent the A. niger proteases of the invention.

XX SQ Sequence 623 AA;

Query Match 32.9%; Score 980.5; DB 23; Length 623; Best Local Similarity 39.6%; Pred. No. 4.8e-77; Matches 216; Conservative 91; Mismatches 178; Indels 61; Gaps 19;

QY 44 TANNVIRKVE--PGAE-----GVCETTPGVKSYSGVVDSPES-----80

DB 32 TPEDLTVIHSEIFPGARISYKQPLGICITTTPTSPSYGIHLPHPTLNLSPGISISQP 91

QY 81 ---HFFWFPEAR--HNPTAPITLWNGPGSDSLIGLFEELPGCHVANSTFDD--YINPH 134

DB 92 YPINTFWFPPSHHNNHTSPUTLWNGPGGSSMIGLFGQNGPCTVNTDSNAYNEW 151

QY 135 SWNEVSNLLFLSPLGVLGVSYSYDTPGNSVTVGVVENSFPAGVQRYPTID--ATLIDT 192

DB 152 SWNEYVDMLYIEQPVOTGFSYDVLNGLTD-----LNETF--LVGTLPSQDVHGVNGT 203

QY 193 TNLAAEAANEILQGLFLSG-----LPSLDSRVQSYD--PSLWTSYGGHYGPAFFNFHFEQNE 247

DB 204 VN--GGRALWVALQWMLGEFSEYVSSVDGNGGGDDRVSIWTSYGGRYGPAYTALTALFORMNE 262

QY 248 RIANGSVN--GVOLNENSLGIINGIIDEAIQAPYYPFAVNTYGIKAVNETVYNYMKFAN 306

DB 263 RIESGEVSTGKKIHLDTLGLHNGCVLLVQVPFPECAVNTYGIQINRTLYDRAMDSW 322

QY 307 QMPNGQDLSTCKQTNRTALAD--YALCAEATNMCNRNVEGYPYAFAGRGVYDIRHPYDD 365

DB 323 SKFGGCRDMIECRDAGE--LGDPLIICEASDYCSREIKSLYNTSGRGYDIAHFTPD 380

QY 366 PTPEPSYNYKFLAKSDVMDATGVNINNTQSNNDVYVAFQQTGDFWPE-----NFTEDLEEI 420

DB 381 AALVPYEVGFLNRPWQKALGPVNYTMSSEAVNSFASTGDI--PRNDRPMIGDILGYL 438

QY 421 LALPVRVSLTYGDADTYCNWPGGQAVSLAANYSAQAQFRSAGYTPPLKXNGVEY--GETRE 478

DB 439 LDSGVKVMYVGRDYACPRWGGEDVSLVVEYDAEKFRAGYAEVQTKS--SYVGGVLVRQ 497

QY 479 YGNFSFTRVYVAGHEVPYQPIASLQLFNRTIFWDIAEGOKKIW--PSYKINGTATATH 536

DB 498 YGNFSFTRVYVQAGHEVPYQPIETAYEIFNRAQFNWDIATGGISLEQNSQYGTGEPSSFWH 557

QY 537 TQSSVP 542

DB 558 IKNEVP 563

RESULT 4

ABR38859 standard; Protein; 554 AA.

XX

AC ABR38859;

XX 24-APR-2003 (first entry)

DT

XX

DE

XX A. niger serine carboxypeptidase polypeptide #2.

XX

KW Protease; fungal infection; aspergillosis; food; tanning; detergent;

KW protein solubility; viscosity; taste; texture; nutritional value;

KW EC3.4.16.6.

OS Aspergillus niger.

XX

XX WO200268623-A2.

XX

XX 06-SEP-2002.

XX

XX 22-FEB-2002; 2002WO-EP01984.

XX

XX 23-FEB-2001; 2001EP-0200657.

XX 23-FEB-2001; 2001EP-0200658.

XX 23-FEB-2001; 2001EP-0200660.

XX 26-FEB-2001; 2001EP-0200706.

XX 26-FEB-2001; 2001EP-0200707.

XX 26-FEB-2001; 2001EP-0200708.

XX 26-FEB-2001; 2001EP-0200719.

XX 28-MAR-2001; 2001EP-0000075.

XX 28-MAR-2001; 2001EP-0000078.

XX 28-MAR-2001; 2001EP-0000080.

XX 28-MAR-2001; 2001EP-0000087.

XX 28-MAR-2001; 2001EP-0000088.

XX 21-MAY-2001; 2001EP-0000156.

XX 21-MAY-2001; 2001EP-0000159.

XX 21-MAY-2001; 2001EP-0000160.

XX 21-MAY-2001; 2001EP-0000162.

XX 21-MAY-2001; 2001EP-0000165.

XX 21-MAY-2001; 2001EP-0000166.

XX 21-MAY-2001; 2001EP-0000168.

XX 21-MAY-2001; 2001EP-0000240.

XX 21-JUN-2001; 2001EP-0000242.

XX 21-JUN-2001; 2001EP-0000244.

XX 21-JUN-2001; 2001EP-0000246.

XX 12-JUL-2001; 2001EP-0000280.

XX 12-JUL-2001; 2001EP-0000285.

XX 30-JUL-2001; 2001EP-0000327.

XX 30-JUL-2001; 2001EP-0000341.

XX 02-AUG-2001; 2001EP-0000342.

XX 02-AUG-2001; 2001EP-0000343.

XX 02-AUG-2001; 2001EP-0000344.

XX 09-AUG-2001; 2001EP-0000357.

XX 16-AUG-2001; 2001EP-0000374.

XX 16-AUG-2001; 2001EP-0000377.

XX 20-SEP-2001; 2001EP-0000478.

XX 20-SEP-2001; 2001EP-0000483.

XX 22-OCT-2001; 2001EP-0000552.

XX 22-OCT-2001; 2001EP-0000553.

XX 22-OCT-2001; 2001EP-0000554.

XX 22-OCT-2001; 2001EP-0000556.

XX 22-OCT-2001; 2001EP-0000557.

XX 22-OCT-2001; 2001EP-0000558.

XX 15-NOV-2001; 2001EP-0000464.

XX 21-DEC-2001; 2001EP-0005117.

XX

FA (STAM) DSM NV.

XX

XX Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;

PI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;

PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;

PI Stiebler J, Albang R;

XX

DR WPI; 2002-723203/78.

DR P-PSDB; ABZ78226, ABZ78283.

XX

PT Novel isolated protease polypeptide useful in laboratory, clinical,
 PT pharmaceutical, chemical, diagnostic, personal care and industrial
 XX applications -
 PS Claim 13; Page 366-369; 394pp; English.
 CC
 CC The invention relates to a novel isolated protease polypeptide. A
 CC polypeptide or polynucleotide of the invention is useful for diagnosing a
 CC fungal infection such as aspergillosis, or as a query sequence to perform
 CC a search against public databases. A polypeptide of the invention is
 CC useful in a selected number of industrial or pharmaceutical processes, in
 CC laboratory or clinical processes, in food industry (baking, brewing,
 CC cheese manufacturing, meat tenderising), in tanning industry and in the
 CC manufacture of biological detergents. A polypeptide may also be useful
 CC for improving protein solubility, extraction yields, viscosity or taste,
 CC texture, nutritional value, minimising of antigenicity or
 CC anti-nutritional factors, colour or functionality as well as processing
 CC aspects like filterability of the proteinaceous raw material. The
 CC sequences shown in ABR38313-ABR38869 represent the A. niger proteases of
 CC the invention.
 XX
 SQ Sequence 554 AA;
 Query Match 31.6%; Score 941; DB 23; Length 554;
 Best Local Similarity 38.5%; Pred. No. 1.2e-73;
 Matches 209; Conservative 72; Mismatches 172; Indels 90; Gaps 13;
 QY 31 QLPKPTGVKTLTA--NNVTIRKEPGAECVCTPGVKSIGYVDTSPEHFFWFFE 88
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 88
 17 QFPEPEGITVLKSLHENVITISKEP--GICETTFGVRSIGYVHLPPASTSFFWFFE 73
 QY 89 ARHNPEPATITLWNGPFGSDSLGLFEELGPCHVNS--TFDDYINPHSWNEVSNLLFSLQ 147
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 147
 74 ARKDPNAPLAIWLNGPFGSSSLMLLEELGPCSIASDSKTTVLNPSWNEVSNLLFSLQ 133
 QY 148 PLGVGFSYSYDVGSIINPVTGVVNSFAGVQGYRPTIDAILDTTNLAEEAAWEILQGF 207
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 207
 134 PTQVGFSDYPTNGTLT-----ANGTAAFAAHLWHFAQTW 168
 QY 208 LSGPLSL---DSRVQSKDFSLWTSYGGHYGPAFFNHFYQNERIANGSV-NGVQ-LNFN 262
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 262
 169 FFEFFHYKPNDRV-----SLWAESYGGHYGPGIFRFFQOQNDKIAEGTAEDGAQYLHLD 223
 QY 263 SLGIINGIIDEALQAPYPEFVANNVYGIKAVNETVYNYMKFANQWPNQCDLISCTKQT 322
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 322
 224 TLGIVNGLMDVQEEAYITWP-----YNNVRLAP-----SSF 256
 QY 323 NRTALADYALCAEATNCRD-----NV-----EGP--YFAPAGRGVYDIR 360
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 360
 257 NSRGFRDQALACEALKERDSGLPHSGKXISEICGGLALEWGDGPITYYHTFNRGWYDIA 316
 QY 361 HPYDDPTPSYNNKFLAKDSVMDAIGVNIYNTQSNNDVYVAFQOQDGFVKNPFIEDLEI 420
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 420
 317 HPKNDPFAKMLGYLTQESVLAALGVNPNFTSSSAVATQFKTFDVIHVGFLDAIGL 376
 QY 421 LALPVRSVLYGDADYICNWFQGVQAVSLAANYSQAFRSAGYTPLKVNGVEYGETREYG 480
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 480
 377 LDGSKVHMVYGDYACNVWGEKASLAVPYSRITEFADTGYSPLLPDGTSGMTRQLG 436
 QY 481 NFSFTRVYAGHEVYQYQIASLQFNRTIFGWDIAEGQKKIWPSTKNGTATATHTQSS 540
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 540
 437 NYSFTRVFOAGHEVSYQPVAAIEYFMRAFNFKNKDIPTGLLAVDDEBFQSVGPKDTHIKNI 496
 QY 541 VPL 543
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 496
 497 PPI 499
 RESULT 5
 ABR38817
 ID ABR38817 standard; Protein; 526 AA.
 XX
 AC ABR38817;
 PI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
 PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
 PI Stiebler J, Albarg R;
 XX WPI; 2002-723203/78.
 DR P-PSDB; ABZ78184, ABZ78241.
 XX Novel isolated protease polypeptide useful in laboratory, clinical,
 PT

XX
 DT 24-APR-2003 (first entry)
 XX
 DE A. niger serine carboxypeptidase polypeptide #1.
 XX
 KW Protease; fungal infection; aspergillosis; food; tanning; detergent;
 KW protein solubility; viscosity; taste; texture; nutritional value;
 KW EC3.4.16.6.
 XX
 OS Aspergillus niger.
 XX
 FN WO200268623-A2.
 XX
 PD 06-SEP-2002.
 XX
 XX 22-FEB-2002; 2002WO-EP01984.
 PF 23-FEB-2001; 2001EP-0200657.
 PR 23-FEB-2001; 2001EP-0200658.
 PR 23-FEB-2001; 2001EP-0200660.
 PR 26-FEB-2001; 2001EP-0200706.
 PR 26-FEB-2001; 2001EP-0200707.
 PR 26-FEB-2001; 2001EP-0200708.
 PR 26-FEB-2001; 2001EP-0200719.
 PR 28-MAR-2001; 2001EP-0000075.
 PR 28-MAR-2001; 2001EP-0000078.
 PR 28-MAR-2001; 2001EP-0000080.
 PR 28-MAR-2001; 2001EP-0000087.
 PR 21-MAY-2001; 2001EP-0000088.
 PR 21-MAY-2001; 2001EP-0000156.
 PR 21-MAY-2001; 2001EP-0000159.
 PR 21-MAY-2001; 2001EP-0000160.
 PR 21-MAY-2001; 2001EP-0000162.
 PR 21-MAY-2001; 2001EP-0000165.
 PR 21-MAY-2001; 2001EP-0000166.
 PR 21-MAY-2001; 2001EP-0000168.
 PR 21-JUN-2001; 2001EP-0000240.
 PR 21-JUN-2001; 2001EP-0000242.
 PR 21-JUN-2001; 2001EP-0000244.
 PR 21-JUN-2001; 2001EP-0000246.
 PR 12-JUL-2001; 2001EP-0000280.
 PR 12-JUL-2001; 2001EP-0000285.
 PR 30-JUL-2001; 2001EP-0000323.
 PR 30-JUL-2001; 2001EP-0000327.
 PR 02-AUG-2001; 2001EP-0000341.
 PR 02-AUG-2001; 2001EP-0000342.
 PR 02-AUG-2001; 2001EP-0000343.
 PR 02-AUG-2001; 2001EP-0000344.
 PR 09-AUG-2001; 2001EP-0000357.
 PR 16-AUG-2001; 2001EP-0000374.
 PR 16-AUG-2001; 2001EP-0000377.
 PR 20-SEP-2001; 2001EP-0000478.
 PR 20-SEP-2001; 2001EP-0000483.
 PR 22-OCT-2001; 2001EP-0000552.
 PR 22-OCT-2001; 2001EP-0000553.
 PR 22-OCT-2001; 2001EP-0000554.
 PR 22-OCT-2001; 2001EP-0000556.
 PR 22-OCT-2001; 2001EP-0000557.
 PR 22-OCT-2001; 2001EP-0000558.
 PR 15-NOV-2001; 2001EP-0004464.
 PR 21-DEC-2001; 2001EP-0005117.
 XX (STAM) DSM NV.
 PA Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
 XX Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
 PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
 PI Stiebler J, Albarg R;
 XX WPI; 2002-723203/78.
 DR P-PSDB; ABZ78184, ABZ78241.
 XX Novel isolated protease polypeptide useful in laboratory, clinical,
 PT

PT pharmaceutical, chemical, diagnostic, personal care and industrial
 XX applications
 XX
 XX Claim 13; Page 271-274; 394pp; English.
 XX
 XX The invention relates to a novel isolated protease polypeptide. A
 CC polypeptide or polynucleotide of the invention is useful for diagnosing a
 CC fungal infection such as aspergillosis, or as a query sequence to perform
 CC a search against public databases. A polypeptide of the invention is
 CC useful in a selected number of industrial or pharmaceutical processes, in
 CC laboratory or clinical processes, in food industry (baking, brewing,
 CC cheese manufacturing, meat tenderising), in tanning industry and in the
 CC manufacture of biological detergents. A polypeptide may also be useful
 CC for improving protein solubility, extraction yields, viscosity or taste,
 CC texture, nutritional value, minimising of antigenicity or
 CC anti-nutritional factors, colour or functionality as well as processing
 CC aspects like filterability of the proteinaceous raw material. The
 CC sequences shown in ABR3813-ABR38869 represent the A. niger proteases of
 XX the invention.
 XX
 XX Sequence 526 AA;
 XX
 XX Query Match 31.1%; Score 925; DB 23; Length 526;
 XX Best Local Similarity 38.0%; Pred. No. 2.8e-72;
 XX Matches 221; Conservative 78; Mismatches 192; Indels 90; Gaps 15;
 XX
 XX 4 YFSLVPLVAVASWALPGSTPASVGRPLPKNPTGKTLTANNVTIRYKPCAEGVCET 63
 XX 2 YSLWVAALVA---ALP-----VSRQFVAPPTDLPTKGYLDIPVRKQV-PTGICET 51
 XX 64 TPGVSYSGVDTSPESHFFPFPEARH-NPETAPITLWNG-----GPGSDSLGLFEE 117
 XX 52 DRSVKFSGYDVAEHEHIFWFPEARNQDTEAPLTVWINGMSDPGPGSSMIGLFOE 111
 XX 118 LGECHVNSFTDDYINPHSNVNSNLLFLSQPLGVGFSYSDTVGSGINPTGVVNSSPAG 177
 XX 112 HGPCGIDANGSVYNNPYSWNASNMLYIDQPVTGFSYS-----IPVPGYVDSSTNG 164
 XX 178 VQGRYPTIDATLDTTNLAARAAWEILQGLSLGPLSDRSVQSKDFSLMSTESYGGHVGPA 237
 XX 165 FMGAPP-----QYRETFHTTESYGGHGPV 191
 XX 238 FNFHFEQNRRIANGSVNGVQLNPNLSGIINGIIDEAIOAPYYPEFAV--NNTYGIKAVN 295
 XX 192 FNEYIEQNAHLQPGA---KKIQLSVMWNGWYDPIIQOAYYNFTVPGNTYDILPEN 248
 XX 296 ETVYNMKFANQMPNGQDILSTCKQNRALADYALCAEATNMCRDNVSGPYAFAGRG 355
 XX 249 KSISLMMNLYGFGNCLDQLYDC-----AARGIDEICSTADDFCANEVENVDIYSGRD 303
 XX 356 VYDIRHPYDPTPSYNNKFLAKDSVMDAIGVNIYTSQNNVDVYAFQQTG-DFVWPENFI 414
 XX 304 EYDFRELTPDPFFVEFYVDYNLAKASVQAAGAYINYNESNNAVLAFSSGTGDLNMTI 363
 XX 415 EDLEEILALPVRSYLIYGDADYICNWFQGGQAVSAAANYSOAAQFRSAGYPLKV-NGVEY 473
 XX 364 QDVGKLLKQGVTVVMYAGDADYNCNMLGGEAVSLQV---KAANFSSAGYINIVTSDGVTH 420
 XX 474 GETREYNFSPTRVYEGHEVYVYOPTASLOLNFETIFGWDIAFGQKKIMPSTYKNTGT-- 531
 XX 421 GQVRQAGQAFVRVYEGSHEVFFYQPLALEMFERRVIGGKDVATGKPISSSLQIVGTPK 480
 XX 532 -----ATATH-TQSSVPLPPTATSMSSVGMVA 555
 XX 481 SYREGNSTIOWEVLDSLATYNTTINAPNPNVSRRLKRMGPA 521
 XX
 XX RESULT 6
 XX ABG93049
 XX ID ABG93049 standard; Protein; 508 AA.
 XX
 XX AC ABG93049;
 XX
 XX

DT 21-NOV-2002 (first entry)
 XX
 XX S. cerevisiae BAX-associated protein fragment SEQ ID 56.
 XX
 XX Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
 XX vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
 XX apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
 XX neurodegeneration; cell death.
 XX
 XX Saccharomyces cerevisiae.
 XX
 XX WO200264766-A2.
 XX
 XX 22-AUG-2002.
 XX
 XX 21-DEC-2001; 2001WO-EP15398.
 XX
 XX 22-DEC-2000; 2000EP-0870318.
 XX
 XX 04-JAN-2001; 2001EP-0870002.
 XX
 XX 09-JAN-2001; 2001EP-0870003.
 XX
 XX (JANC) JANSSEN PHARM NV.
 XX
 XX Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
 XX
 XX WPI; 2002-667002/71.
 XX
 XX N-PSDB; ABQ76315.
 XX
 XX New isolated nucleic acid representing a synthetic Bax gene, useful as
 XX medicament for treating, preventing and/or alleviating yeast or fungal
 XX infections or proliferative disorders, or for preventing apoptosis in
 XX certain diseases
 XX
 XX Claim 36; Figure 1; 344pp; English.
 XX
 XX This invention describes a novel nucleic acid representing a synthetic
 XX Bax gene. The Bax gene of the invention is useful for identifying
 XX Bax-resistant yeast or fungi, identifying, or obtaining and identifying
 XX Candida spp. sequences that are differentially expressed in a pathway
 XX eventually leading to programmed cell death. The products of the
 XX inhibitor sequences of Bax-induced cell death. The products of the
 XX invention have cytostatic, fungicide; immunosuppressive, virucide and
 XX vasotropic activity and can be used in vaccines or for gene therapy. The
 XX isolated nucleic acids, polypeptides, pharmaceutical compositions,
 XX antisense molecules and antibodies are useful as medicaments or in
 XX preparing a medicament for treating, preventing and/or alleviating
 XX diseases associated with yeast or fungi or proliferative disorders, such
 XX as cancer, or for preventing apoptosis in certain diseases. The compounds
 XX or polypeptides, or the genetically modified organism are useful for
 XX preparing a medicament for modifying the endogenic flora of humans and
 XX other mammals. The vaccine is useful for immunising against yeast or
 XX fungal infections. Apoptosis-related diseases include autoimmune disease,
 XX ischaemia, diseases related with viral infections or neurodegenerations.
 XX This sequence represents a polypeptide associated with the Bax gene
 XX described in the disclosure of the invention.
 XX
 XX Sequence 508 AA;
 XX
 XX Query Match 21.3%; Score 636; DB 23; Length 508;
 XX Best Local Similarity 32.6%; Pred. No. 6.5e-47;
 XX Matches 167; Conservative 70; Mismatches 197; Indels 78; Gaps 14;
 XX
 XX 19 LPGSTPASVGRQL----BKNPTGVKTLTANNVTIRYKPCAEGVCETTPGVKSYSGYV 74
 XX 43 LPQNTQTLKDLRLNHHDDPLFTFPISSVDDTDYSRLRLTVDPKSLGI----DTVQWSSGYM 98
 XX 75 DTPSESTHFFWFPEARHNPETAPITLWNGPGSDSLGLFEEGPGCHVNSFTDDYINPH 134
 XX 99 DYKSKHFFWFPEARNDPANDPILWNGPGSCSFTGLFELGPGSSIGADMKPIHNPY 158
 XX 135 SWNEVSNLLFLSQPLGVGFSYSDTVGSGINPTGVVNSSPAGVQGRPTIDATLIDTIN 194
 XX 159 SWNNNASMIFLEQLGVGFSYGD-----EKVSSTK 188

QY	284	AVNNTYGIKAVNETVYMKFANOMPNGCQDLISTCKQTNRTALADYALCAEATNMCRDN	343
Db	305	ACGGGYPVAVLDESSCQMD--NALPR-CQSMIESCYSESA-----WVCVPASYICNNA	356
QY	344	VEGPIYAFAGRGVDIRHEYDDP-----TPPSYNNKFLAKSDVMDAIGVNNIN-YTQSNND	397
Db	357	LLAPYQR-TGQNVYDVRGKCEDSSNLCYSAMGYSDYLKPEVIEAVGAENVGYDSCNPD	415
QY	398	VYVAFQQTGDFWVP--NFIEDLEELALPVRVSLIY-GDADYICNWFQGGQAVSLAANYSQ	454
Db	416	INRNFLFHGDMWPKYHRLVPGLE--QIPV---LIYAGDADFICNWLGNKAWTEALEWPG	470
QY	455	AAQPSAGVTPLKV-----NGVEYGETREYGNFSTRVYEAGHEVPYQPIASLQLFNRT	509
Db	471	QAEYASAELEDLIVVDNEHTGKKIQVKSNGFTFMRLYGGGHVPMQDPESSELEFFNRW	530
QY	510	IFG-W 513	
Db	531	LGGEW 535	
RESULT 8			
ID	AAR96737	standard; Protein; 556 AA.	
AC	AAR96737;		
DT	14-AUG-1996	(first entry)	
DE	A. niger	Bo-1 carboxypeptidase Y.	
KW	Carboxypeptidase Y; CPY; ascomycete; Deuteromycete; host cell;		
KM	protease deficiency.		
OS	Aspergillus niger strain Bo-1.		
PN	WO9609397-Al.		
PD	28-MAR-1996.		
PF	19-SEP-1995;	95WO-US11945.	
PR	20-SEP-1994;	94US-0309341.	
PA	(NOVO) NOVO NORDISK BIOTECH INC.		
PI	Thompson SA, Yaver DS;		
DR	WPI; 1996-188458/19.		
DR	N-PSDB; AAT28283.		
PT	Nucleic acid construct encoding a filamentous ascomycete or		
PT	deuteromycete carboxy:peptidase Y - useful to produce host cells		
PT	modified to produce reduced amounts of carboxy:peptidase		
PS	Claim 4; Fig 1A-E; 46pp; English.		
CC	Carboxypeptidase Y (AAR96737), a vacuolar protease, is the product of		
CC	the CPY gene (AAT28283) of Aspergillus niger strain Bo-1. Creation		
CC	of CPY-deficient Aspergillus strains, e.g. by cloning a selectable		
CC	marker into the CPY gene, provides suitable host strains for prodn.		
CC	of heterologous protein.		
XX	Sequence	556 AA;	
QY	Query Match	21.1%; Score 630; DB 17; Length 556;	
Db	Best Local Similarity	31.8%; Pred. No. 2.5e-46;	
	Matches 154; Conservative	81; Mismatches 174; Indels 76; Gaps 17;	
QY	45	ANNVTIRYKEPCAGVCECTTPGKSGYGVDTSP-ESHTFFWFFFEARHNPTAPITLWLN	103
Db	131	AYDLRVKKTDPGSLGI---DPGVKQYTGVLDDNENDKHLFYWFFESRNDPNDPVVLWLN	187

PS	Claim 5; Page 23-25; 46pp; English.	
CC	Carboxypeptidase Y (AAR96738), a vacuolar protease, is the product of the CPY gene (AAT28284) of <i>Aspergillus niger</i> strain SFAG 2. Creation of CPY-deficient <i>Aspergillus</i> strains, e.g. by cloning a selectable marker into the CPY gene, provides suitable host strains for prodn. of heterologous protein.	
XX		
SQ	Sequence 557 AA;	
	Query Match	
	Best Local Similarity 31.8%; Score 620; DB 17; Length 557;	
	Matches 15; Conservative 80; Mismatches 175; Indels 76; Gaps 18;	
QY	45 ANNVTIRKPEPGAGVCETTPGVKSYGYVDTSF-ESHTRFWPEAEARNHETAPITILWLN 103	
DB	132 AYDLRVKKTDPSSLGI--DPGVKYQYIGYLDNDENKHLFWFFESRNDPENDFWVLWLN 188	
QY	104 GPGGSDSLGLFBEELGCHVNSTFDDYINPHSWNEVSNLLFLSOPLCWGFSGYSYDVTGSI 163	
DB	189 GPGGCSLLTGLFMELGPPSSINKKIQPVYNDYAWNSNASVIFLDQPVNVGYSYNSA---- 244	
QY	164 NPVTGVVENSFPAGVQGRYPTIDATLIDTNNLAABAAWEILQGLSLGSLDSRVQSKDF 223	
DB	245 -----VSDTVAAGKQVYALLTLFFKQFP----EYAKQDF 274	
QY	224 SLWTESYGGHYGPAFFNFHFEQNERIANGSVGVQLNFNSLIGTINGIIDEAICQAPYYPEF 283	
DB	275 HIAGESYAGHYIVPFASEILSHKR-----NINLQSVLIGNGLTDLGLTEYFYRPM 325	
QY	284 AVNNTYGIKAVNETVYNYMKFANQMPNGCODLISTCKQTNRATADYALCAEATNCRDN 343	
DB	326 ACDGQ-GYPAVLDE-GSCQAMDNALPR-CQSMIESCYSSSA-----WVCVPASITCNNA 377	
QY	344 VEGPYAFAGRGYVDIRHPYDDP-----TPPSYNNKFLAKDSYMDAIGVNNI-YTQSNND 397	
DB	378 LIAPYQR-TGQNYVDVRGKCEDSSNLCYSAMGVSDYLNKTEVIEAVGAENVGYDSCNFD 436	
QY	398 VYVAFQQTGFVWP--NFIEDLEEILALPVRVSLIY-GDADYICNWFQGOAVSLAANYSQ 454	
DB	437 INENFLFHGDWMPYHRLVPGLE--QIPV---LIYAGDADFICNLGNKAWTEALEWPG 491	
QY	455 AAQFRSAGYTPLKV-----NGVEYGETREYGNFSTRVYEAGHEVPVYQPIASLQLFNET 509	
DB	492 QAQYASAKLEDLVVNEHKGKKGQVKSHGNFTFNLRYGGGHVVPMDQFESSLEFFNRW 551	
QY	510 IFG-W 513	
DB	552 LGGEW 556	
RESULT 10		
AAR48059		
ID	AAR48059 standard; Protein; 491 AA.	
XX		
AC	AAR48059;	
XX		
DT	25-MAR-2003 (updated)	
DT	20-JUL-1994 (first entry)	
XX		
DE	Sequence of protease C encoded by gene Kl.PRC1.	
XX		
KW	Protease; yeast; proteolysis.	
OS	Kluyveromyces lactis.	
XX		
FN	W09400579-Al.	
XX		
XX	06-JAN-1994.	
PF	23-JUN-1993; 93WO-FR00623.	
XX		
XX	25-JUN-1992; 92FR-0007785.	

Db 273 ---ASGFIINNRSEFELASVLNGLTNGHTDPLIQGSY-----

Qy 304 FANQMENGCG----ODLIST--CKQWRTALADYALCAEATNMCRD-----NVEGPY-- 348

Db 306 ----KPMGCCEGGYPVLTTEQCDDMER----DYPRCAKLTKLCYSFOSALTCPAQYVC 357

Qy 349 -----YAFAGRGVYDIRHPYDDP-----TPPSYYNKFLAKDSVMDAIGV-NIN-YTQS 394

Db 358 DSRILFPQYAQTGLNPFYDIRKDCAEQGCCGYVEMDYLDLYLNLDDYVKAEVGASNIDIFTSC 417

Qy 395 NNDVYFAFOQTGDPEWFNFIEDLEEIIALPYRVSLIYGDAFYICNWFGGQAVSLAANSQ 454

Db 418 DDVTFRNFLDGDEMKP-FCQYVAELLDNNVFVIYAGDKDYICMWLGNLAWNVELEISD 476

Qy 455 AAQFSRAGYTPLKVNGYE-VGETREYGNFSFTRVVYEAHEVPYQPIASLQLFNRTIFG 512

Db 477 SEHFAPLPQLWKQDGKAAGEVKNHKHFTFLRIYDAGHMVFDFDPENALSMTWTWVOG 535

RESULT 12

ID ABR38845 standard; Protein; 481 AA.

XX AC ABR38845;

XX DT 24-APR-2003 (first entry)

XX DE A. niger carboxypeptidase Y polypeptide #3.

XX KW Protease; fungal infection; aspergilliosis; food; tanning; detergent;

KW protein solubility; viscosity; taste; texture; nutritional value;

KW EC3.4.16.5.

XX CS Aspergillus niger.

XX PN WO2000268623-A2.

XX PD 06-SEP-2002.

XX PF 22-FEB-2002; 2002WO-EP01984.

XX PR 23-FEB-2001; 2001EP-0200657.

PR 23-FEB-2001; 2001EP-0200658.

PR 23-FEB-2001; 2001EP-0200660.

PR 26-FEB-2001; 2001EP-0200706.

PR 26-FEB-2001; 2001EP-0200707.

PR 26-FEB-2001; 2001EP-0200708.

PR 26-FEB-2001; 2001EP-0200719.

PR 28-MAR-2001; 2001EP-0000075.

PR 28-MAR-2001; 2001EP-0000078.

PR 28-MAR-2001; 2001EP-0000080.

PR 28-MAR-2001; 2001EP-0000087.

PR 28-MAR-2001; 2001EP-0000088.

PR 21-MAY-2001; 2001EP-0000156.

PR 21-MAY-2001; 2001EP-0000159.

PR 21-MAY-2001; 2001EP-0000160.

PR 21-MAY-2001; 2001EP-0000162.

PR 21-MAY-2001; 2001EP-0000165.

PR 21-MAY-2001; 2001EP-0000166.

PR 21-MAY-2001; 2001EP-0000168.

PR 21-JUN-2001; 2001EP-0000240.

PR 21-JUN-2001; 2001EP-0000242.

PR 21-JUN-2001; 2001EP-0000244.

PR 21-JUN-2001; 2001EP-0000246.

PR 12-JUL-2001; 2001EP-0000280.

PR 12-JUL-2001; 2001EP-0000285.

PR 30-JUL-2001; 2001EP-0000323.

PR 30-JUL-2001; 2001EP-0000327.

PR 02-AUG-2001; 2001EP-0000341.

PR 02-AUG-2001; 2001EP-0000342.

PR 02-AUG-2001; 2001EP-0000343.

PR 02-AUG-2001; 2001EP-0000344.

C. albicans BAX-associated protein fragment SEQ ID 520.

Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide; vasotropic; vaccine; gene therapy; proliferative disorder; cancer; apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia; neurodegeneration; cell death.

Candida albicans.
WO200264766-A2.
22-AUG-2002.

21-DEC-2001; 2001WO-EP15398.
22-DEC-2000; 2000EP-0870318.
04-JAN-2001; 2001EP-0870002.
09-JAN-2001; 2001EP-0870003.
(JANC) JANSSEN PHARM NV.

Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
WPI : 2002-667002/71.
N-PSDB; ABQ76547.

New isolated nucleic acid representing a synthetic Bax gene, useful as medicament for treating, preventing and/or alleviating yeast or fungal infections or proliferative disorders, or for preventing apoptosis in certain diseases -
Claim 36; Figure 2; 344pp; English.

This invention describes a novel nucleic acid representing a synthetic Bax gene. The Bax gene of the invention is useful for identifying Bax-resistant yeast or fungi, identifying, or obtaining and identifying Candida spp. sequences that are differentially expressed in a pathway eventually leading to programmed cell death or identifying inhibitors or inhibitor sequences of Bax-induced cell death. The products of the invention have cytostatic, fungicide; immunosuppressive, virucide and vasotropic activity and can be used in vaccines or for gene therapy. The isolated nucleic acids, polypeptides, pharmaceutical compositions, antisense molecules and antibodies are useful as medicaments or in preparing a medicament for treating, preventing and/or alleviating diseases associated with yeast or fungi or proliferative disorders, such as cancer, or for preventing apoptosis in certain diseases. The compounds or polypeptides , or the genetically modified organism are useful for preparing a medicament for modifying the endogenous flora of humans and other mammals. The vaccine is useful for immunising against yeast or fungal infections. Apoptosis-related diseases include autoimmune disease, ischaemia, diseases related with viral infections or neurodegenerations. This sequence represents a polypeptide associated with the Bax gene described in the disclosure of the invention.

XX
XX Sequence 550 AA;

SQ

Query Match 18.2%; Score 541.5; DB 23; Length 550;
Best Local Similarity 31.3%; Pred. No. 1.5e-38;
Matches 150; Conservative 63; Mismatches 157; Indels 109; Gaps 17;

QY 67 VKSYGYVD-TSPSHTFFWFEEAFHNDETPATILWLNGGPGSDSLIGLFEELGPCVNS 125
|||:|||: |||:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:
DB 133 VKQVTGLDIDSLDKHLYFWFFESRNDPKNDPILWLNGGPGCSSTGLFELGPSSINK 192
|||:|||: |||:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:

QY 126 TFDYINPHSNVEYSNLFLSQPLGVGFYSYSTDVGSINPVTVGVVENSFAGVGQRYPTI 185
|||:|||: |||:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:

DB 193 TLHPVPYNPSWNNSNAIFLDQPVGVGYSY-----TGGE----- 227
|||:|||: |||:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:

QY 186 DATLDTITNLAAEAAWEILQGFLSGPLDSRGVSQDFSLMTESYGHYGPAFFNHFYEQ 245
|||:|||: |||:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:

DB 228 ----VKNTLTAAKQVVVFLELFQKEP----QLFNKTHIAGESVAGHYIPAF----- 272
|||:|||: |||:|||||:|:|:|||||:|:|:|||||:|:|:|||||:|:

[illegible]

us-09-712-338-2.rag

Mon Nov 24 13:41:08 2003

PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0142154.	PR	16-SEP-1999;	99US-0154039.
PR	02-JUL-1999;	99US-0142055.	PR	20-SEP-1999;	99US-0154779.
PR	06-JUL-1999;	99US-0142390.	PR	22-SEP-1999;	99US-0155139.
PR	08-JUL-1999;	99US-0142803.	PR	23-SEP-1999;	99US-0155486.
PR	09-JUL-1999;	99US-0142920.	PR	24-SEP-1999;	99US-0155659.
PR	12-JUL-1999;	99US-0142977.	PR	28-SEP-1999;	99US-0156458.
PR	13-JUL-1999;	99US-0143342.	PR	29-SEP-1999;	99US-0156596.
PR	14-JUL-1999;	99US-0143624.	PR	04-OCT-1999;	99US-0157117.
PR	15-JUL-1999;	99US-0144005.	PR	05-OCT-1999;	99US-0157753.
PR	16-JUL-1999;	99US-0144085.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144086.	PR	07-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144325.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145919.	PR	25-OCT-1999;	99US-0161406.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146389.	PR	26-OCT-1999;	99US-0161361.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161920.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161992.
PR	04-AUG-1999;	99US-0147302.	PR	28-OCT-1999;	99US-0161993.
PR	05-AUG-1999;	99US-0147192.	PR	29-OCT-1999;	99US-0162142.

Query Match

12.9%; Score 383; DB 21; Length 482;

Best Local Similarity 26.2%; Pred. No. le-24;	
Matches 135; Conservative 68; Mismatches 202; Indels 110; Gaps 1	
Qy	33 PKNPTGVKTLTANNVTIRYKE-----PGAEGVCETTPG-----VKSYGYVDTSPESH 82
Db	32 PSPKRGVSSGDTSHFNVIQRESVSPKDKDLIQLPGQSDVTFKQYGYVAVNKPAGR 91
Qy	83 F--TWFFARINPETAITLWNGPGSDSLGLFEELGPHCHVNSTFDD--YINPHSWNEV 139
Db	92 FLYYYFVETIKPGNTTPLVWFNGPGGCSLLGGAFKELGPFVHSDGKTLFENPYSWNE 151
Qy	140 SNLLFLSPLGVGSYSVDVGSINPVTGVVENSFAGVQGRYPTIDATLIDTTNLAAPA 199
Db	152 ANVLFLTEPVTGFSYS-----NSPINGKQGDAT-----AED 184
Qy	200 AWEILQGLSLGSLDSRVQSKDFSLWTESYGGHYGPAFFNHFFYQNERIANGSVNGVL 259
Db	185 NYMELVNWLEFPPEYKGR-----DIYAGQSVAGHYVPOLAQIILHRN-----NQTLI 232
Qy	260 NFNSLGIINGHIDEAIQAPYYPEFAVNNTYGIKAVNETVYNYMKFANQMPNGCDLISTC 319
Db	233 NLRGILIGNPSLREIQDDFGYKMF--SHGLIS--QQQMDYNYKF-----CTD--SDL 280
Qy	320 KQTNRTALADYALCAEATNMCRDNVEGPPYYAGRGVYDIRHPY----- 363
Db	281 YDWKCHLASOKIEAQKTHL-----DIYNIAPLCNLSTLSSEPKKCTTI 325
Qy	364 --DDETPPSYNYKFLAKDSVMDAIGVN-----INVTQSNNDVYYAFQOTGDFVWFNFIED 416
Db	326 MKADPCSGNYLKAYLNKEVQEAIHANTKPIYEWTSNTKLLMEWNKDRYV--SLTPI 383
Qy	417 LEEILALPVRVSLYIGDADYICNWFQGOAVSLAANYSOAOFRSAGYTPLKVGVEYGET 476
Db	384 LQELMGKGVVRWLYNGVDVLIPTSTLAVVKTKMLTVVKEWR-----PWFTGGHVGGFT 438
Qy	477 REY--GNFSFRVYAGHEVPYYPQIASLQLNRTI 510
Db	439 EDYKGNLTFVTYKGAHGSVPTDQPIHALNIFTSI 473
RESULT 14	
AAG30064	
ID	AAG30064 standard; Protein; 502 AA.
XX	AAG30064;
AC	
DT	17-OCT-2000 (first entry)
DX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 35876.
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.

PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145088.
PR	21-JUL-1999;	99US-0145088.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0147038.
PR	03-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159283.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0144352.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 12.9%; Score 383; DB 21; Length 502;		
Best Local Similarity 26.2%; Pred. No. 1.le-24; Indels 110; Gaps 18;		
Matches 135; Conservative 68; Mismatches 202;		
QY	33	PKNPTGVKTLTTANNVTIRYKE-----PGAEGVCETTPG-----VKSYSGYDTSPEST 82
Db	52	PSFKGVSSSGDTSHEINVQRESVSPKDKLIQLPGQSDVTFKQGGYVAVNKPAGR 111
QY	83	F-FWFFEARHNPETAPITLWNGFGSDSLGLFEELGCPHNSTFDP-YINPHSWNEV 139
Db	112	FLYVVFETIKPGNTITPLVWENGPGCSSLGAFKELGPFVHSDGKTLFRNPSWNE 171
QY	140	SNLLFLSQPLGVGFSYSDVDSINPVTGVWNSFAGVQGRYPTIDATLIDTTLAAEA 199
Db	172	ANVLFLETPVGTGFSYS-----NSPINGQGDYAT-----AED 204
QY	200	AMEILQGLSLPSLDSRVQSKDFSLWTSYGHYGPAPFNHFYEQNERIANGSVNGVOL 259
Db	205	NYMFLVNWLERFPEYKGR---DIYTAGOSYAGHYVPLQAQILHFN-----NOTLI 252
QY	260	NENSLGIINGIDEAIQAPYEPFVANNYTGKAVNETVYNMKFANQMPNGQDLISCTC 319
Db	253	NLRGILIGNPSLNREIQDDDFGYKMF--SHGLIS-OQQMDNKNKF-----CTD--SDL 300
QY	320	KQTNRTALADYALCAEATNMCRDNVGRVPYAFAGRGVYDIRHPY-----363
Db	301	YDWDKCHLASQKTEAOKTHL-----DIYNIYAPLCNLSTLSSEPKCTTI 345
QY	364	--DDPTPPSYNFKLAKDSVMDAIGN-----INYTSNNNDVYAFQQTGDFVWPNFIED 416
Db	346	MKADPCSGNYLKAYLNKEVOEALHANTTKIPYEWTSCTKLWENNEKDRYV--SLTPI 403
QY	417	LEBILALPVRSLLIYGDADYICNWFQGOAVSLAANSQAQPSAGYVTLKVNQVEYGET 476
Db	404	LOELMGKGVRLYNGEDVDIVIPFTSTLAVTKTMNLTVVKEWR-----PWFTGCHVGFFT 458
QY	477	REY-GNFSFTRVVEAGHEVPYQPIASLQLENRTI 510
Db	459	EDYKGNLFTVTVKGAGHSVTDQPIHALNIFTSEFI 493
RESULT 15		
ABR36843		
ID	ABR36843	standard; Protein; 455 AA.
XX	ABR36843;	
XX	ABR36843;	
DT	24-APR-2003	(first entry)
XX		

DE A. niger carboxypeptidase Y polypeptide #2.
 XX Protease; fungal infection; aspergillosis; food; tanning; detergent;
 KW protein solubility; viscosity; taste; texture; nutritional value;
 KW EC3.4.16.5.
 XX Aspergillus niger.
 XX WO200268623-A2.
 XX 06-SEP-2002.
 XX 22-FEB-2002; 2002WO-EP01984.
 XX 23-FEB-2001; 2001EP-0200657.
 PR 23-FEB-2001; 2001EP-0200658.
 PR 23-FEB-2001; 2001EP-0200660.
 PR 26-FEB-2001; 2001EP-0200706.
 PR 26-FEB-2001; 2001EP-0200707.
 PR 26-FEB-2001; 2001EP-0200708.
 PR 26-FEB-2001; 2001EP-0200719.
 PR 28-MAR-2001; 2001EP-0000075.
 PR 28-MAR-2001; 2001EP-0000078.
 PR 28-MAR-2001; 2001EP-0000080.
 PR 28-MAR-2001; 2001EP-0000087.
 PR 28-MAR-2001; 2001EP-0000088.
 PR 21-MAY-2001; 2001EP-0000156.
 PR 21-MAY-2001; 2001EP-0000159.
 PR 21-MAY-2001; 2001EP-0000160.
 PR 21-MAY-2001; 2001EP-0000162.
 PR 21-MAY-2001; 2001EP-0000165.
 PR 21-MAY-2001; 2001EP-0000166.
 PR 21-MAY-2001; 2001EP-0000168.
 PR 21-JUN-2001; 2001EP-0000240.
 PR 21-JUN-2001; 2001EP-0000242.
 PR 21-JUN-2001; 2001EP-0000244.
 PR 21-JUN-2001; 2001EP-0000246.
 PR 12-JUL-2001; 2001EP-0000280.
 PR 12-JUL-2001; 2001EP-0000285.
 PR 30-JUL-2001; 2001EP-0000323.
 PR 02-AUG-2001; 2001EP-0000341.
 PR 02-AUG-2001; 2001EP-0000342.
 PR 02-AUG-2001; 2001EP-0000343.
 PR 02-AUG-2001; 2001EP-0000344.
 PR 09-AUG-2001; 2001EP-0000357.
 PR 16-AUG-2001; 2001EP-0000374.
 PR 16-AUG-2001; 2001EP-0000377.
 PR 20-SEP-2001; 2001EP-0000478.
 PR 20-SEP-2001; 2001EP-0000483.
 PR 22-OCT-2001; 2001EP-0000552.
 PR 22-OCT-2001; 2001EP-0000553.
 PR 22-OCT-2001; 2001EP-0000554.
 PR 22-OCT-2001; 2001EP-0000556.
 PR 22-OCT-2001; 2001EP-0000557.
 PR 22-OCT-2001; 2001EP-0000558.
 PR 15-NOV-2001; 2001EP-0000464.
 PR 21-DEC-2001; 2001EP-0005117.
 XX (STAM) DSM NV.
 XX Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
 PI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
 PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
 PI Stiebler J, Albarg R;
 XX WPI: 2002-723203/78.
 DR P-PSDB; ABZ78210, ABZ78267.
 DR Novel isolated protease polypeptide useful in laboratory, clinical,
 XX pharmaceutical, chemical, diagnostic, personal care and industrial
 PT applications -

PS Claim 13; Page 329-331; 394pp; English.
 XX

CC The invention relates to a novel isolated protease polypeptide. A
 CC polypeptide or polynucleotide of the invention is useful for diagnosing a
 CC fungal infection such as aspergillosis, or as a query sequence to perform
 CC a search against public databases. A polypeptide of the invention is
 CC useful in a selected number of industrial or pharmaceutical processes, in
 CC laboratory or clinical processes, in food industry (baking, brewing,
 CC cheese manufacturing, meat tenderising), in tanning industry and in the
 CC manufacture of biological detergents. A polypeptide may also be useful
 CC for improving protein solubility, extraction yields, viscosity or taste,
 CC texture, nutritional value, minimising of antigenicity or
 CC anti-nutritional factors, colour or functionality as well as processing
 CC aspects like filterability of the proteinaceous raw material. The
 CC sequences shown in ABR38313-ABR38869 represent the A. niger proteases of
 CC the invention.
 XX

SQ Sequence 455 AA;

Query Match 12.3%; Score 365; DB 23; Length 455;
 Best Local Similarity 27.3%; Pred. No. 3.5e-23;
 Matches 135; Conservative 68; Mismatches 197; Indels 94; Gaps 22;

QY	21	GSTPASVGRRLPKNPTGVKTLTANNVTIRYKPEAGVCEITPGVKSXSGYVDTSPES	80
DB	25	GADPFTVFRS--PHSPA-----FSIRIQEN-DSICDARS--PQFTGLMDIGPK-	68
QY	81	HTFFWFPEARHNPETAPITLWNGPGSDSLIGLIERLGPCHVN--STFDDYINPHSWNE	138
DB	69	HLFFWYFESQNDFFHDPDLTLMWTGGPCDSSWIGLFEVFGPCRFNEFGNGTDH-NFWATK	127
QY	139	VSNLLFSLQPLGVGFYSYSDTVDSINPVTGVNSSFAGVQGRYPTIDATLITTNLAEE	198
DB	128	NSSLFVDPQVDVGFYSID-----EGYE-----LPHDSREAAV	160
QY	199	AAWEILLQGLSLGLPSLDSRVQSKDFSLWTEYSGHYGPAPFNHFYQNERIANGSVNGVQ	258
DB	161	DMRFLRLFLIS---EIFPHKQFLPVHLGSGESYAGRYPIYLATQILEQNELYKOSP----	213
QY	259	LNFNLSGIINGIIDEAIAIQAPYYPEFAVNNITYGIKA--VNETVYNYMKFANQMENGCDLI	316
DB	214	IPLKSLVGVNGFMSPKDATFGYWEITCTTNSGVPSPIFNETRCDIM--AANMEH-CMDLY	270
QY	317	STCKQTNRTALADYALCAEATNCRDNVEGPPYAFAGRGVYDIRHPYDDFTPP-----	369
DB	271	DICIQH-----SDPAICHAAQSVYDSVVGMLAKLLR-----MTVTAPCEIDEMC	317
QY	370	----SYNKFLLAKDSVMDAIGVNNINYTQSN---NOVYVAFQQTGDFVWPNFIEDLEI---	420
DB	318	YIEAALIERYLNSPSWEALSPQQVTEYKFVATSVIDAFQASAD----GMVSSSKQIAF	373
QY	421	-LALPVRVSLIYGDADYICNWFEGGQAVSLAANYSOAAQFERSAGYTPKLVNGVEYGE-TRE	478
DB	374	LLANNVDFLAYQGNLDLACNTAGNLRWANSLSMKGQTEFTAKPLLPWEIQ-VSVGEGTDE	432
QY	479	YGNFSFTRVYEAGH	492
DB	433	TSRFAFVTVDNAGH	446

Search completed: November 21, 2003, 17:59:37
 Job time : 47 secs

This Page Blank (uspto)

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 18:01:36 ; Search time 39 Seconds
 (without alignments)
 3672.287 Million cell updates/sec

Title: US-09-712-338-2
 Perfect score: 555
 Sequence: 1 MRGYFSLVPLVAASWALP.....HTQSSVPLTATSMSSVGMA 555

Scoring table: OLIGO
 Gapex 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Query		Description		
Result No.	Score	Match	Length	DB	ID			
1	454	81.8	555	3	Q96VZ9	Q96VZ9 aspergillus		
2	10	1.8	460	3	Q12569	Q12569 absidia zyc		
3	9	1.6	67	2	Q937G9	Q937G9 klebsiella		
4	9	1.6	147	2	Q936V3	Q936V3 pseudomonas		
5	9	1.6	405	2	O66024	O66024 pseudomonas		
6	9	1.6	405	2	P71195	P71195 enterobacte		
7	9	1.6	405	2	Q91UM7	Q91UM7 rhizobium m		
8	9	1.6	405	2	O56451	O56451 xanthomonas		
9	9	1.6	405	2	Q9F3W1	Q9F3W1 pseudomonas		
10	9	1.6	468	10	Q9SHU3	Q9SHU3 arabidopsis		
11	9	1.6	469	10	Q9FH06	Q9FH06 arabidopsis		
12	9	1.6	473	10	Q9FH05	Q9FH05 arabidopsis		
13	9	1.6	479	10	Q9MAR8	Q9MAR8 arabidopsis		
14	9	1.6	495	2	Q48384	Q48384 klebsiella		
15	9	1.6	2105	5	Q17679	Q17679 caenorhabdi		
16	8	1.4	116	15	Q75691	Q75691 human immun		

17	8	1.4	166	17	Q59236	Q59236 pyrococcus
18	8	1.4	263	2	Q9ZF75	Q9ZF75 burkholderi
19	8	1.4	264	10	Q9LXY6	Q9LXY6 arabidopsis
20	8	1.4	264	10	Q8GVD2	Q8GVD2 zea mays lm
21	8	1.4	270	16	Q9RRC8	Q9RRC8 deinococcus
22	8	1.4	329	17	Q8TURS	Q8TURS methanopyru
23	8	1.4	362	10	Q9XE81	Q9XE81 sorghum bic
24	8	1.4	425	10	O82229	O82229 arabidopsis
25	8	1.4	425	10	O65568	O65568 arabidopsis
26	8	1.4	434	11	Q8BKV9	Q8BKV9 mus musculu
27	8	1.4	437	10	Q9FWG1	Q9FWG1 oryza sativ
28	8	1.4	437	11	Q9CZK1	Q9CZK1 mus musculu
29	8	1.4	437	11	Q9DBG5	Q9DBG5 mus musculu
30	8	1.4	447	10	Q8S8K6	Q8S8K6 arabidopsis
31	8	1.4	452	10	Q9ZQO0	Q9ZQO0 arabidopsis
32	8	1.4	456	10	Q9SV78	Q9SV78 arabidopsis
33	8	1.4	456	10	Q9C7E2	Q9C7E2 arabidopsis
34	8	1.4	458	10	O22803	O22803 arabidopsis
35	8	1.4	459	10	Q8RZS0	Q8RZS0 oryza sativ
36	8	1.4	459	10	Q9SFB5	Q9SFB5 arabidopsis
37	8	1.4	461	10	Q93Y09	Q93Y09 arabidopsis
38	8	1.4	465	10	O04084	O04084 arabidopsis
39	8	1.4	465	10	Q94K84	Q94K84 arabidopsis
40	8	1.4	465	10	Q8VY01	Q8VY01 arabidopsis
41	8	1.4	465	10	Q9M099	Q9M099 arabidopsis
42	8	1.4	470	10	O22732	O22732 arabidopsis
43	8	1.4	472	10	Q9LSM9	Q9LSM9 arabidopsis
44	8	1.4	474	11	Q9D2D1	Q9D2D1 mus musculu
45	8	1.4	479	10	Q949Q7	Q949Q7 arabidopsis

ALIGNMENTS

RESULT 1

Q96VZ9 PRELIMINARY; PRT; 555 AA.

AC Q96VZ9, 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Carboxypeptidase S1 precursor (EC 3.4.16.6).

GN Cpl.

OS Aspergillus oryzae.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OX NCBI_TaxID=5062;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=TK3;

RA van den Broek P.;

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF394242; AAK77166.1; -

DR InterPro; IPR001563; Serine_carpept.

DR Pfam; PF00450; serine_carpept; 1.

DR PRINTS; PR00724; CRBOXYPTASEC.

DR ProDom; PD001189; Serine_carpept; 1.

DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.

KW Carboxypeptidase; Hydrolase; Signal.

FT SIGNAL 1 18

ST SEQUENCE 555 AA; 61168 MW; 456B63B0CEB55222 CRC64;

Query Match 81.8%; Score 454; DB 3; Length 555;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 554; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRGYFSLVPLVAASWALPGSTPASVGRQLPKNPTGVKTLTTANNVTIRKEPAGGV 60
 Db 1 MRGYFSLVPLVAASWALPGSTPASVGRQLPKNPTGVKTLTTANNVTIRKEPAGGV 60

QY 61 CETTGVKSYSGYVDTSPESHFTFFFFFAHNPETAPITLWNGFGSDSLGLFEELGP 120
 Db 61 CETTGVKSYSGYVDTSPESHFTFFFFFAHNPETAPITLWNGFGSDSLGLFEELGP 120

Db 430 RVEAGHEVP 439

QY 121 CHVNSTFDYINPHSWNEVSNLLFLSQPLGVGFSYSYDVTGDSINPVTGWENSSPAGVQG 180
DB 121 CHVNSTFDYINPHSWNEVSNLLFLSQPLGVGFSYNDVDSINPVTGWENSSPAGVQG 180
QY 181 RYPTIDATLIDTTLAAEAAEILQGLSLDSRVQSKDFSLWTSYGGHYGPAPFN 240
DB 181 RYPTIDATLIDTTLAAEAAEILQGLSLDSRVQSKDFSLWTSYGGHYGPAPFN 240
QY 241 HYEONERIANGVNGVQVNFNSLGIINGIDEAICAPYYPFAVNNYVGIKAVNETVYN 300
DB 241 HYEONERIANGVNGVQVNFNSLGIINGIDEAICAPYYPFAVNNYVGIKAVNETVYN 300
QY 301 YMKFANQMPNGQDILISTCKQNRNTALADYALCAEATNMCRDNVEGPPYAFAGRGVYDIR 360
DB 301 YMKFANQMPNGQDILISTCKQNRNTALADYALCAEATNMCRDNVEGPPYAFAGRGVYDIR 360
QY 361 HPYDDPTPPSYNKKFLAKDSVMDAIGVNNYNTQSNNDVYVAFQOTGDFVWPNFIEDLEE 420
DB 361 HPYDDPTPPSYNKKFLAKDSVMDAIGVNNYNTQSNNDVYVAFQOTGDFVWPNFIEDLEE 420
QY 421 LALPVRVSLIYGDADVICNWFEGQAVSLAANYSOAAQFRSAGYTPPLKXNGVVEYGETREYG 480
DB 421 LALPVRVSLIYGDADVICNWFEGQAVSLAANYSOAAQFRSAGYTPPLKXNGVVEYGETREYG 480
QY 481 NFSFTRVYAGHEVPYQYQIASQLFNRTIFGWDIAEGQKKIWPSTKNTGTATATHTQSS 540
DB 481 NFSFTRVYAGHEVPYQYQIASQLFNRTIFGWDIAEGQKKIWPSTKNTGTATATHTQSS 540
QY 541 VPLPTATSMSSVGMA 555
DB 541 VPLPTATSMSSVGMA 555

RESULT 2

Q12569 ID Q12569 PRELIMINARY; PRT; 460 AA.
AC Q12569
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Prepro-carboxypeptidase Z (EC 3.4.16.5).
GN SPCZ.
OS Absidia zychnae.
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Absidia.
OC NCBI_TaxID=36079;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRIC 1199;
RX MEDLINE=95308538; PubMed=7788719;
RA Lee B.R., Takeuchi M., Kobayashi Y.,
RT "Molecular cloning and sequence analysis of the scpZ gene encoding the
RL serine carboxypeptidase of Absidia zychnae."
RL Curr. Genet. 27:159-165(1995).
DR EMBL; D16519; BAA03966.1; --
DR HSSP; P00729; 1YSC.
DR MEROPS; S10.010; --
DR Interpro; IPR001563; Serine_carbpept.
DR Interpro; IPR000379; Serine_carbpept.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXPTASEC.
DR ProDom; PD001189; Serine_carbpept; 1.
DR Carboxypeptidase; Hydrolase.
FT CHAIN 52 460 MATURE CARBOXYPEPTIDASE Z.
SQ SEQUENCE 460 AA; 50825 MW; CE3A4107645488B6 CRC64;
Query Match 1.8%; Score 10; DB 3; Length 460;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 486 RVEAGHEVP 495
DB 486 RVEAGHEVP 495

Db 430 RVEAGHEVP 439

RESULT 3
Q937G9 ID Q937G9 PRELIMINARY; PRT; 67 AA.
AC Q937G9
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative transposition protein (fragment).
GN TNIQ.
OS Klebsiella sp. LS13-39.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=143776;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LS13-39; TRANSPOSON=Tn5056;
RX MEDLINE=21604134; PubMed=11763242;
RA Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
RA Yurieva O.V., Nikiforov V.G.;
RT "Mercury resistance transposons of Gram-negative environmental
RT bacteria and their classification."
RL Res. Microbiol. 152:811-822(2001).
DR EMBL; AJ027770; CAC82957.1; --
FT NON_TER 1
SQ SEQUENCE 67 AA; 7547 MW; 3BECC0A1149C62D9 CRC64;
Query Match 1.6%; Score 9; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARNHPETA 96

DB 3 EARNHPETA 11

RESULT 4

Q936V3 ID Q936V3 PRELIMINARY; PRT; 147 AA.
AC Q936V3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative transposition protein tniQ (fragment).
GN TNIQ.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BW20; TRANSPOSON=Tn5053V4;
RX MEDLINE=21604134; PubMed=11763242;
RA Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
RA Yurieva O.V., Nikiforov V.G.;
RT "Mercury resistance transposons of Gram-negative environmental
RT bacteria and their classification."
RL Res. Microbiol. 152:811-822(2001).
DR EMBL; AJ296313; CAC82974.1; --
FT NON_TER 1
SQ SEQUENCE 147 AA; 16500 MW; 24636FA7A1489ABD CRC64;
Query Match 1.6%; Score 9; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 88 EARNHPETA 96
DB 83 EARNHPETA 91

```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 EARNPETA 96
Db 341 EARNPETA 349

RESULT 7
Q91UM7
ID Q91UM7 PRELIMINARY; PRT; 405 AA.
AC Q91UM7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE ThiQ protein.
GN ThiQ.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSB102.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=mercury resistance transposon Tn5718;
RA Schneiker S., Keller M., Droegge M., Lanka E., Puhler A.,
RA Selbitschka W.;
RT "The genetic organization and evolution of the broad-host-range
RT mercury resistance plasmid pSB102 isolated from a microbial population
RT residing in the rhizosphere of alfalfa.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ304453; CAC79210.1;
KW Plasmid.
SQ SEQUENCE 405 AA; 45711 MW; 2396F2C89A2DE7B2 CRC64;

Query Match 1.6%; Score 9; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 EARNPETA 96
Db 341 EARNPETA 349

RESULT 8
Q56451
ID Q56451 PRELIMINARY; PRT; 405 AA.
AC Q56451;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE Mercury resistant transposon TNS053 (Central Asia).
GN TNS053.
OS Xanthomonas sp.
OG Plasmid RPI.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=29446;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93253772; PubMed=8387603;
RA Kholodii G.Y.A., Yurieva O.V., Lomovskaya O.L., Gorlenko Zh.M.,
RA Mindlin S.Z., Nikiforov V.G.;
RT "TNS053, a mercury resistance transposon with integron's ends.";
RL J. Mol. Biol. 230:1103-1107(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96130850; PubMed=8594337;
RA Kholodii G.Y., Mindlin S.Z., Bass I.A., Yurieva O.V., Minakhina S.V.,
RA Nikiforov V.G.;
RT "Four genes, two ends, and a res region are involved in transposition
RT of TNS053: a paradigm for a novel family of transposons carrying either
RT a mer operon or an integron.";
RL Mol. Microbiol. 17:1189-1200(1995).

```

```

[3]
RN  SEQUENCE FROM N.A.
RP  Kholodii G.Y.;
RA  Russ. J. Genet. 31:1447-1451 (1995).
RL  EMBL; L40595; AAA98331.1; -.
RW  Plasmid.
KW  Plasmid.
SQ  SEQUENCE 405 AA; 45830 MW; 89880020E32684C3 CRC64;

Query Match 1.6%; Score 9; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARNHPETA 96
Db 341 EARNHPETA 349

RESULT 9
ID Q9F3W1 PRELIMINARY; PRT; 405 AA.
AC Q9F3W1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE TniQ protein.
DE TniQ.
GN TniQ.
OS Pseudomonas sp. ED23-33.
OC Bacteria; Proteobacteria.
OC NCBI_TaxID=83781;
RN [1]
RP SEQUENCE FROM N.A.
RA Minakshina S., Minakhin L., Kholodii G., Mindlin S., Gorlenko Z.H.,
RA Yuriyeva O., Nikiforov V.;
RT "Molecular inventory of transposons from environmental bacteria:
RT epidemic dissemination of Tn21-, Tn5041-, and Tn5053-related
RT transposons."
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y17897; CAC34709.1; -.
SQ SEQUENCE 405 AA; 45706 MW; 033A913AF44E43EE CRC64;

Query Match 1.6%; Score 9; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARNHPETA 96
Db 341 EARNHPETA 349

RESULT 10
ID Q9SHU3 PRELIMINARY; PRT; 468 AA.
AC Q9SHU3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative serine carboxypeptidase II.
GN A2G12480.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.I., Barnstead M.B., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,

[3]
RN  SEQUENCE FROM N.A.
RP  Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA  Salzberg S.L., Fraser C.M., Venter J.C.;
RT  "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RL  thaliana."
RL  Nature 402:761-768 (1999).
RW  Nature 402:761-768 (1999).
KW  Nature 402:761-768 (1999).
SQ  SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007268; AAD28662.1; -.
DR HSSP; P08819; 1WHT.
DR MEROPS; S10.005; -.
DR InterPro; IPR002000; Lamp.
DR InterPro; IPR001563; Serine carboxypept.
DR Pfam; PF00450; serine carboxypept; 1.
DR PRINTS; PR00724; CRBOXYPASEC.
DR ProDom; PD001189; Serine carboxypept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
DR PROSITE; PS00310; LAMP_1; 1.
KW Carboxypeptidase.
KW SEQUENCE 468 AA; 52836 MW; 47DD2055AB9AE1D7 CRC64;

Query Match 1.6%; Score 9; DB 10; Length 468;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWLNCGPG 107
Db 77 TLWLNCGPG 85

RESULT 11
ID Q9FH06 PRELIMINARY; PRT; 469 AA.
AC Q9FH06;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Serine carboxypeptidase-II like.
DE Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:31-63 (2000).
DR EMBL; AB023032; BAB10196.1; -.
DR HSSP; P08819; 1WHT.
DR MEROPS; S10.005; -.
DR InterPro; IPR001563; Serine carboxypept.
DR Pfam; PF00450; serine carboxypept; 1.
DR PRINTS; PR00724; CRBOXYPASEC.
DR ProDom; PD001189; Serine carboxypept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase.
KW SEQUENCE 469 AA; 52541 MW; 90357E27B9B64527 CRC64;

Query Match 1.6%; Score 9; DB 10; Length 469;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWLNCGPG 107
Db 74 TLWLNCGPG 82

```

```

RESULT 12
Q9FH05 PRELIMINARY; PRT; 473 AA.
AC Q9FH05;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Serine carboxypeptidase II-like (AF5942240/K5J14.4).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler E., Chao Q., Chin C., Chiou J., Choi E., Gonzalez A.,
RA Howling B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharsy N., Pham P., Sakano H., Shinn P.,
RA Toriumi M., Vayshberg M., Yu G., Ecker J., Theologis A., Davis R.W.;
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC006423; AAF63101.1; -.
DR HSSP; P08819; 1WHT.
DR MEROPS; S10.005; -.
DR InterPro; IPR001563; Serine carboxypeptidase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00450; serine_carboxypeptidase.
DR PRINTS; PR00724; CRBOXYPASEC.
DR ProDom; PD001189; Serine carboxypeptidase.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase.
SQ SEQUENCE 473 AA; 52874 MW; 714E0F29BC251590 CRC64;

Query Match 1.6%; Score 9; DB 10; Length 473;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 TLWINGGPG 107
Db 78 TLWINGGPG 86

RESULT 13
Q9MAR8 PRELIMINARY; PRT; 479 AA.
AC Q9MAR8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Putative serine carboxypeptidases.
GN F28H19.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler E., Chao Q., Chin C., Chiou J., Choi E., Gonzalez A.,
RA Howling B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharsy N., Pham P., Sakano H., Shinn P.,
RA Toriumi M., Vayshberg M., Yu G., Ecker J., Theologis A., Davis R.W.;
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC006423; AAF63101.1; -.
DR HSSP; P08819; 1WHT.
DR MEROPS; S10.005; -.
DR InterPro; IPR001563; Serine carboxypeptidase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00450; serine_carboxypeptidase.
DR PRINTS; PR00724; CRBOXYPASEC.
DR ProDom; PD001189; Serine carboxypeptidase.
DR PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
KW Carboxypeptidase.
SQ SEQUENCE 479 AA; 53922 MW; 87A36BB1C46E1080 CRC64;

Query Match 1.6%; Score 9; DB 10; Length 479;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 TLWINGGPG 107
Db 83 TLWINGGPG 91

RESULT 14
Q48384 PRELIMINARY; PRT; 495 AA.
AC Q48384;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE ORF6 protein.
OS Klebsiella aerogenes.
OG Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
NCBI_TaxID=28451;
RN [1]
RP SEQUENCE FROM N.A.
RA Sundstroem L.;
RA Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=94252994; PubMed=8195081;
RA Radstrom P., Skold O., Swedberg G., Flensburg J., Roy P.H.,
RA Sundstrom L.;
RA "Transposon Tn5090 of plasmid R751, which carries an integron, is
RT related to Tn7, Mu, and the retroelements.";
RL J. Bacteriol. 176:3257-3268 (1994).
KW Plasmid.
SQ SEQUENCE 495 AA; 55107 MW; A36EF0E3F3B70B13 CRC64;

Query Match 1.6%; Score 9; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 3.8;

```

Mon Nov 24 13:41:08 2003

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARNHETA 96
 |||||
 DB 431 EARNHETA 439

RESULT 15

Q17679 PRELIMINARY; PRT; 2105 AA.
 AC Q17679;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE Y16B4A.2 protein.
 GN Y16B4A.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thomas K.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z70203; CAA94110.1; -.
 DR EMBL; AL023825; CAA94110.1; JOINED.
 DR EMBL; Z93339; CAA94110.1; JOINED.
 DR EMBL; Z93339; CAB07544.1; -.
 DR EMBL; AL023825; CAB07544.1; JOINED.
 DR EMBL; Z70203; CAB07544.1; JOINED.
 DR EMBL; AL023825; CRA19443.1; -.
 DR EMBL; Z93339; CRA19443.1; JOINED.
 DR EMBL; Z70203; CAA19443.1; JOINED.
 DR HSP; P10619; LIVY.
 DR WormPep; Y16B4A.2; CE21374.
 DR InterPro; IPR001563; Serine_carbpept.
 DR InterPro; IPR000379; Ser_estrs_site.
 DR Pfam; PF00450; serine_carbpept; 4.
 DR PRINTS; PR00724; CRBOXYPTASEC.
 DR ProDom; PD001189; Serine_carbpept; 6.
 DR PROSITE; PS00560; CARBOXYEPT_SER_HIS; 4.
 DR PROSITE; PS00131; CARBOXYEPT_SER_SER; 2.
 SQ SEQUENCE 2105 AA; 234335 MW; 139CB6E96D75B3FB CRC64;

Query Match 1.6%; Score 9; DB 5; Length 2105;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPGS 108
 |||||
 DB 549 LWLNGGPGS 557

Search completed: November 21, 2003, 18:05:16
 Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 18:01:06 ; Search time 17 Seconds
(without alignments)
1535.284 Million cell updates/sec

Title: US-09-712-338-2

Perfect score: 555

Sequence: 1 MRGYEFLSVLPVLAASWALP.....HTQSSVPLPTATSMSSVGMMA 555

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	1.8	423	1	CPS1_PENJA
2	9	1.6	508	1	YBY9_YEAST
3	8	1.4	36	1	PSBY_ODOSI
4	8	1.4	423	1	CBP2_WHEAT
5	8	1.4	436	1	CP22_HORVU
6	8	1.4	470	1	YSS2_CAEEL
7	8	1.4	474	1	PRTP_MOUSE
8	8	1.4	476	1	CBP2_HORVU
9	8	1.4	480	1	PRTP_HUMAN
10	8	1.4	499	1	CBP1_HORVU
11	8	1.4	510	1	CBP1_ORYSA
12	8	1.4	516	1	CP23_HORVU
13	8	1.4	523	1	CBPY_PICPA
14	8	1.4	532	1	CBPY_YEAST
15	8	1.4	542	1	CBPY_CANAL
16	8	1.4	574	1	YPP3_CAEEL
17	8	1.4	1002	1	CBPY_SCHPO
18	8	1.4	4349	1	FAT2_HUMAN
19	7	1.3	25	1	VCX8_ODOSI
20	7	1.3	110	1	RLA4_YEAST
21	7	1.3	117	1	ARR2_ECOLI
22	7	1.3	158	1	GML_HUMAN
23	7	1.3	216	1	OL1F_HUMAN
24	7	1.3	216	1	OL1I_HUMAN
25	7	1.3	225	1	SOC1_MOUSE
26	7	1.3	250	1	PSD8_CAEEL
27	7	1.3	290	1	SUHB_MYCTO
28	7	1.3	297	1	VGLG_HRSV3
29	7	1.3	298	1	VGLG_HRSVA
30	7	1.3	298	1	VGLG_HRSVL
31	7	1.3	315	1	O3A1_HUMAN
32	7	1.3	315	1	O3A3_HUMAN
33	7	1.3	317	1	PLC_LISMO

34 7 1.3 321 1 STILL_DROME P92189 drosophila
35 7 1.3 334 1 ADD_VIBCH Q9KNI7 vibrio chol
36 7 1.3 341 1 GLK_RHILO Q98DM2 rhizobium l
37 7 1.3 344 1 DFO4_STRPN Q97SC7 streptococc
38 7 1.3 346 1 PA6A_MOUSE O9Z101 mus musculus
39 7 1.3 353 1 Y740_AQUAE O6E948 aquifex ae
40 7 1.3 354 1 DP42_RHIME Q9ZXH8 rhizobium m
41 7 1.3 356 1 DPO4_STAAM Q99SZ1 staphylococ
42 7 1.3 356 1 E13L_TOBAC P58964 staphylococ
43 7 1.3 356 1 MURG_YERPE P52399 nicotiana t
44 7 1.3 356 1 MURG_YERPE O8ZIE9 versinia p
45 7 1.3 357 1 DP42_AGRTS Q8UJK7 agrobacteri

ALIGNMENTS

RESULT 1

CPS1_PENJA
ID - CPS1_PENJA STANDARD; PRT; 423 AA.
AC P34946;
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carboxypeptidase S1 (EC 3.4.16.6)
OS Penicillium janthinellum (Penicillium vitale).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
OX NCBI_TaxID=5079;
RN [1]
RP SEQUENCE.
RX MEDLINE=94039747; PubMed=8224168;
RA Svendsen I., Hofmann T., Endrizzi J., Remington S.J., Breddam K.;
RT "The primary structure of carboxypeptidase S1 from Penicillium janthinellum";
RL FEBS Lett. 333:39-43(1993).
CC -!- CATALYTIC ACTIVITY: Preferential release of a C-terminal arginine or lysine residue.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
DR PIR; S38953; S38953.
DR HSSP; P00729; LYSC.
DR MEROPS; S10.008; -.
DR InterPro; IPR000379; Ser_estrs site.
DR InterPro; IPR001563; Serine_carbpept.
DR Pfam; PF00450; serine_carbpept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR ProDom; PD001189; Serine_carbpept; 1.
DR PROSITE; PS00131; CARBOXYPEPT_SER_HIS; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydrolase; Carboxypeptidase; Glycoprotein.
FT ACT_SITE 143 143 POTENTIAL.
FT ACT_SITE 340 340 BY SIMILARITY.
FT BINDING 343 343 BY SIMILARITY.
FT BINDING 398 398 SUBSTRATE.
FT DISULFID 8 68
FT DISULFID 55 300
FT DISULFID 223 246
FT DISULFID 230 239
FT CARBOHYD 200 200
SQ SEQUENCE 423 AA; 46496 MW; 55F3333608B12729 CRC64;
N-LINKED (GLCNAC...).

Query Match 1.8%; Score 10; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 0.046; 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0;

QY 227 TSYGGHYGCP 236

Db 141 TSYGGHYGCP 150

RESULT 2

YBY9_YEAST

```

ID YB9 YFAST STANDARD; PRT; 508 AA.
AC P38109;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative serine carboxypeptidase in ESR1-IRAI intergenic region
DE (EC 3.4.16.-).
GN YBR139W OR YBR1015.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c;
RX MEDLINE=94378717; PubMed=8091856;
RA Becam A.-M., Cullin C., Grzybowska E., Lacroute F., Nasr F.,
RA Ozier-Kalogeropoulos O., Palucha A., Slonimski P.P., Zagulski M.,
RA Herbert C.J.;
RT "The sequence of 29.7 kb from the right arm of chromosome II reveals
RT 13 complete open reading frames, of which ten correspond to new
RT genes.";
RL Yeast 10:S1-S11(1994).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=95042830; PubMed=7954890;
RA Nasr F., Becam A.-M., Grzybowska E., Zagulski M., Slonimski P.P.,
RA Herbert C.J.;
RT "An analysis of the sequence of part of the right arm of chromosome
RT II of S. cerevisiae reveals new genes encoding an amino-acid permease
RT and a carboxypeptidase.";
RL Curr. Genet. 26:1-7(1994).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X75891; CAA53497.1; -.
CC EMBL; Z36008; CAA85097.1; -.
CC PIR; S46008; S46008.
CC HSSP; P00729; 1CPY.
CC MEROPS; S10 UPW; -.
CC SGD; S0000343; YBR139W
CC InterPro; IPR000379; Ser_estr_ssite.
CC InterPro; IPR001563; Serine_carbpept.
CC Pfam; PF00450; serine_carbpept; 1.
CC PRINTS; PR00724; CRBOXVPTASEC.
CC ProDom; PD001189; Serine_carbpept; 1.
CC PROSITE; PS00131; CARBOXYPEPT_SER_HIS; 1.
CC PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hypothetical protein; Hydrolase; Carboxypeptidase.
FT ACT_SITE 219 BY SIMILARITY.
FT ACT_SITE 219 BY SIMILARITY.
FT ACT_SITE 415 BY SIMILARITY.
FT ACT_SITE 474 BY SIMILARITY.
FT ACT_SITE 474 BY SIMILARITY.
SQ SEQUENCE 508 AA; 57639 MW; AAB2806C8EE2ED81 CRC64;

Query Match 1.6%; Score 9; DB 1; Length 508;
Best Local Similarity 100.0%; Pred.No. 0.58;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 QPLGVGFPSY 155
DB 171 QPLGVGFPSY 179

RESULT 3
PSBY ODOSI
ID PSBY_ODOSI STANDARD; PRT; 36 AA.

ID YB9 YFAST STANDARD; PRT; 508 AA.
AC P38109;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative serine carboxypeptidase in ESR1-IRAI intergenic region
DE (EC 3.4.16.-).
GN YBR139W OR YBR1015.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c;
RX MEDLINE=94378717; PubMed=8091856;
RA Becam A.-M., Cullin C., Grzybowska E., Lacroute F., Nasr F.,
RA Ozier-Kalogeropoulos O., Palucha A., Slonimski P.P., Zagulski M.,
RA Herbert C.J.;
RT "The sequence of 29.7 kb from the right arm of chromosome II reveals
RT 13 complete open reading frames, of which ten correspond to new
RT genes.";
RL Yeast 10:S1-S11(1994).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=95042830; PubMed=7954890;
RA Nasr F., Becam A.-M., Grzybowska E., Zagulski M., Slonimski P.P.,
RA Herbert C.J.;
RT "An analysis of the sequence of part of the right arm of chromosome
RT II of S. cerevisiae reveals new genes encoding an amino-acid permease
RT and a carboxypeptidase.";
RL Curr. Genet. 26:1-7(1994).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X75891; CAA53497.1; -.
CC EMBL; Z36008; CAA85097.1; -.
CC PIR; S46008; S46008.
CC HSSP; P00729; 1CPY.
CC MEROPS; S10 UPW; -.
CC SGD; S0000343; YBR139W
CC InterPro; IPR000379; Ser_estr_ssite.
CC InterPro; IPR001563; Serine_carbpept.
CC Pfam; PF00450; serine_carbpept; 1.
CC PRINTS; PR00724; CRBOXVPTASEC.
CC ProDom; PD001189; Serine_carbpept; 1.
CC PROSITE; PS00131; CARBOXYPEPT_SER_HIS; 1.
CC PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hypothetical protein; Hydrolase; Carboxypeptidase.
FT ACT_SITE 219 BY SIMILARITY.
FT ACT_SITE 219 BY SIMILARITY.
FT ACT_SITE 415 BY SIMILARITY.
FT ACT_SITE 474 BY SIMILARITY.
FT ACT_SITE 474 BY SIMILARITY.
SQ SEQUENCE 508 AA; 57639 MW; AAB2806C8EE2ED81 CRC64;

Query Match 1.6%; Score 9; DB 1; Length 508;
Best Local Similarity 100.0%; Pred.No. 0.58;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 QPLGVGFPSY 155
DB 171 QPLGVGFPSY 179

RESULT 3
PSBY ODOSI
ID PSBY_ODOSI STANDARD; PRT; 36 AA.

```

```

AC P49543;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem II protein Y.
DE (PSBY-A OR PSBY-1 OR YCF32.1) AND (PSBY-B OR PSBY-2 OR YCF32.2).
OS Odontella sinensis (Marine centric diatom).
OC Chloroplast.
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Bidulphiophyceae; Eupodiscaceae; Eupodiscaceae; Odontella.
OX NCBI_TaxID=2839;
RN [1]
RP SEQUENCE FROM N.A.
RX Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RT Odontella sinensis.";
RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC -!- FUNCTION: MANGANESE-BINDING POLYPEPTIDE WITH L-ARGININE
CC METABOLIZING ENZYME ACTIVITY. COMPONENT OF THE CORE OF PHOTOSYSTEM
CC II (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -!- SIMILARITY: BELONGS TO THE PSBY FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z67753; CAA91612.1; -.
CC EMBL; Z67753; CAA91666.1; -.
CC PIR; S78239; S78239.
CC KW Photosystem II; Chloroplast; Transmembrane; Thylakoid.
CC FT DOMAIN 1 6 LUMENAL (POTENTIAL).
CC FT TRANSMEM 7 23 POTENTIAL.
CC FT DOMAIN 24 36 STROMAL (POTENTIAL).
CC SEQUENCE 36 AA; 3921 MW; 22989D3E7DAE2B63 CRC64;

Query Match 1.4%; Score 8; DB 1; Length 36;
Best Local Similarity 100.0%; Pred.No. 0.58;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LVAASWAL 19
DB 13 LVAASWAL 20

RESULT 4
CBP2 WHEAT STANDARD; PRT; 423 AA.
AC P08819;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase II chains A and B (EC 3.4.16.6)
DE (Carboxypeptidase D) (CPDW-II) (CP-WII).
GN CBP2.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE.
RX Breddam K., Soerensen S.B., Svendsen I.;
RT "Primary structure and enzymatic properties of carboxypeptidase II
RT from wheat bran.";
RL Carlsberg Res. Commun. 52:297-311(1987).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).
RX MEDLINE=90216664; PubMed=2324089;

```


FT	TURN	158	159
FT	HELIX	160	174
FT	TURN	175	175
FT	TURN	177	178
FT	STRAND	180	187
FT	STRAND	192	192
FT	HELIX	193	205
FT	TURN	206	208
FT	HELIX	212	222
FT	TURN	223	224
FT	HELIX	232	245
FT	TURN	250	251
FT	TURN	253	254
FT	TURN	269	270
FT	HELIX	271	279
FT	TURN	280	280
FT	HELIX	282	288
FT	TURN	289	289
FT	TURN	292	293
FT	STRAND	303	303
FT	HELIX	305	309
FT	TURN	310	310
FT	STRAND	317	317
FT	HELIX	319	327
FT	TURN	328	329
FT	STRAND	331	337
FT	TURN	338	339
FT	HELIX	345	353
FT	TURN	354	356
FT	STRAND	359	368
FT	TURN	369	370
FT	STRAND	371	379
FT	TURN	380	381
FT	STRAND	382	387
FT	TURN	388	389
FT	HELIX	394	397
FT	HELIX	399	411
FT	TURN	412	412
SQ	SEQUENCE	423 AA; 47111 MW; C64815BFF62E085C CRC64;	

Query Match 1.4%; Score 8; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	100	LWLNGGPG	107
DB	57	LWLNGGPG	64

|||||||
|||||||

RESULT 5
CP22_HORVU STANDARD; PRT; 436 AA.
AC P55746;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine carboxypeptidase II-2 precursor (EC 3.4.16.6) (CP-MII.2) (Fragment).
GN CXP;2-2.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
EC STRAIN=cv. Alexis; TISSUE=Grain;
RX MEDLINE=94336715; PubMed=7520177;
RA Dal Began F., Rocher A., Cameron-Mills V., von Wettstein D.;
RT "The expression of serine carboxypeptidases during maturation and
RT germination of the barley grain."
RL Proc. Natl. Acad. Sci. U.S.A. 91:8209-8213 (1994).
RG 1. CAMERON-MILLS V., ROCHER A., DAL BEGAN F., WETTSTEIN D. (1994) EXPRESSION OF SERINE CARBOXYPEPTIDASES DURING MATURATION AND GERMINATION OF THE BARLEY GRAIN. PROC. NATL. ACAD. SCI. U.S.A. 91: 8209-8213.

or lysine residue.
-!- SUBUNIT: CARBOXYPEPTIDASE II IS A DIMER, WHERE EACH MONOMER IS COMPOSED OF TWO CHAINS LINKED BY A DISULFIDE BOND (BY SIMILARITY).
-!- DEVELOPMENTAL STAGE: EXPRESSED IN THE GERMINATING EMBRYO. LOW LEVELS IN THE DEVELOPING ALEURONE AND EMBRYO. ALSO FOUND IN THE ROOTS AND SHOOTS OF THE GROWING SEEDLING.
-!- PTM: THE LINKER PEPTIDE IS ENDOPROTEOLYTICALLY EXCISED DURING ENZYME MATURATION (BY SIMILARITY).
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
EMBL; X78878; CAB59202.1; -.
HSSP; P08819; 1WHT.
InterPro; IPR000379; Ser esters site.
InterPro; IPR001563; Serine carboxypept.
Pfam; PF00450; serine carboxypept; 1.
PRINTS; PR00724; CRBOXYPTASEC.
ProDom; PD001189; Serine carboxypept; 1.
ProSITE; PS00131; CARBOXYPEPT SER SER; 1.
ProSITE; PS00560; CARBOXYPEPT SER HIS; 1.
Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Multigene family.
KW NON_TER 1 1 SERINE CARBOXYPEPTIDASE II-2, CHAIN A.
FT CHAIN <1 256 LINKER PEPTIDE (BY SIMILARITY).
FT PROPEP 257 270 SERINE CARBOXYPEPTIDASE II-2, CHAIN B.
FT CHAIN 271 436 BY SIMILARITY.
FT ACT_SITE 149 149 BY SIMILARITY.
FT ACT_SITE 350 350 BY SIMILARITY.
FT ACT_SITE 403 403 BY SIMILARITY.
FT ACT_SITE 403 403 INTERCHAIN (BY SIMILARITY).
FT DISULFID 56 313 BY SIMILARITY.
FT DISULFID 217 229 BY SIMILARITY.
FT DISULFID 253 281 BY SIMILARITY.
FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 436 AA; 48952 MW; 80F82D97E0C34DC9 CRC64;
Query Match 1.4%; Score 8; DB 1; Length 436;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 100 LWLNGGPG 107
Db 48 LWLNGGPG 55
RESULT 6
YSS2 CAEEL
ID_YSS2 CAEEL STANDARD; PRT; 470 AA.
AC Q09991;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative serine carboxypeptidase K10B2.2 precursor (EC 3.4.16.-).
GN K10B2.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Miller N.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
EMBL; U28730; AAA68259.1; -.
PIR; T16606; T16606.
HSSP; P10619; LIIV.
MEROPS; S10.002; -.
InterPro; K10B2.2; CEO2009.
InterPro; IPR000379; Ser esters site.
InterPro; IPR001563; Serine carboxypept.
Pfam; PF00450; serine carboxypept; 1.
PRINTS; PR00724; CRBOXYPTASEC.
ProDom; PD001189; Serine carboxypept; 2.
ProSITE; PS00131; CARBOXYPEPT SER SER; 1.
ProSITE; PS00560; CARBOXYPEPT SER HIS; 1.
Hypothetical protein; Hydrolase; Carboxypeptidase; Glycoprotein;
KW Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 470 PUTATIVE SERINE CARBOXYPEPTIDASE K10B2.2.
FT ACT_SITE 169 169 BY SIMILARITY.
FT ACT_SITE 380 380 BY SIMILARITY.
FT ACT_SITE 441 441 BY SIMILARITY.
FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 470 AA; 53158 MW; CCC2DACB75EF30FC CRC64;
Query Match 1.4%; Score 8; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 100 LWLNGGPG 107
Db 71 LWLNGGPG 78
RESULT 7
PRTP MOUSE
ID_PRTP MOUSE STANDARD; PRT; 474 AA.
AC P16675; Q8VEF6;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Lysosomal protective protein precursor (EC 3.4.16.5) (Cathepsin A) (Carboxypeptidase C) (protective protein for beta-galactosidase).
GN PPG8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE=90170982; PubMed=2106523;
RA Galjart N.J.; Gillenans N.; Meijer D.; D'Azzo A.;
RT "Mouse protective protein". cDNA cloning, sequence comparison, and expression.";
RL J. Biol. Chem. 265:4678-4684 (1990).
RN 2;
RP SEQUENCE FROM N.A.
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA Klausner R.D.; Collins F.S.; Wagner L.; Sherman C.M.; Schuler G.D.;
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;
RA Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.;
RA Raha S.S.; Loquellano N.A.; Peters G.J.; Abramson R.D.; Mullahy S.J.;
RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Huliyk S.W.;
RA Villalón D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.

FT MOD_RES 314 314 BLOCKED.
 FT VARIANT 351 352 TN -> AT.
 FT CONFLICT 181 181 Y -> R (IN REF. 2).
 SQ SEQUENCE 476 AA; 52625 MW; 18685725B1A6B5E4 CRC64;
 Query Match 1.4%; Score 8; DB 1; Length 476;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 100 LWINGPG 107
 Db 89 LWINGPG 96
 RESULT 9
 PRTP_HUMAN STANDARD; PRT; 480 AA.
 AC P10619; Q96KJ2; Q9BW69;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Lysosomal protective protein precursor (EC 3.4.16.5) (Cathepsin A)
 DE (Carboxypeptidase C) (Protective protein for beta-galactosidase).
 GN PPG8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=88311078; PubMed=3136930;
 RA Galjart N.J., Gillemans N., Harris A., van de Horst G.T.J.,
 RA Verheijen F.W., Galjaard H., D'Azso A.,
 RT "Expression of cDNA encoding the human 'protective protein'
 RT associated with lysosomal beta-galactosidase and neuraminidase:
 RT homology to yeast proteases.";
 RL Cell 54:755-764 (1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharshaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McClay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871 (2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting M., Madan A., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP SEQUENCE OF 29-53 AND 327-351.
 RC TISSUE=Platelet;
 RX MEDLINE=90293074; PubMed=1694176;
 RA Jackman H.L., Tan F., Tamei H., Beuriling-Harbury C., Li X.-Y.,
 RA Skidgel R.A., Erdoes E.G.;
 RT "A peptidase in human platelets that deamidates tachykinins. Probable
 RT identity with the lysosomal 'protective protein'.";
 RL J. Biol. Chem. 265:11265-11272 (1990).
 RN [5]
 RP FUNCTION, AND MUTAGENESIS.
 RX MEDLINE=91117848; PubMed=1907282;
 RA Galjart N.J., Morreau H., Willemsen R., Gillemans N., Bonten E.J.,
 RA D'Azso A.;
 RT "Human lysosomal protective protein has cathepsin A-like activity
 RT distinct from its protective function.";
 RL J. Biol. Chem. 266:14754-14762 (1991).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=96164441; PubMed=8591035;
 RA Rudenko G., Bonten E., D'Azso A., Hol W.G.J.;
 RT "Three-dimensional structure of the human 'protective protein':
 RT structure of the precursor form suggests a complex activation
 RT mechanism.";
 RL Structure 3:1249-1259 (1995).
 RN [7]
 RP VARIANT GALACTOSIALIDOSIS VAL-440.
 RX MEDLINE=92097522; PubMed=1756715;
 RA Zhou X.Y., Galjart N.J., Willemsen R., Gillemans N., Galjaard H.,
 RA D'Azso A.;
 RT "A mutation in a mild form of galactosialidosis impairs dimerization
 RT of the protective protein and renders it unstable.";
 RL EMBO J. 10:4041-4048 (1991).
 RN [8]
 RP VARIANTS GALACTOSIALIDOSIS ARG-49; ARG-65; LEU-90; ASN-249 AND
 RP CYS-395.
 RX MEDLINE=93293970; PubMed=8514852;
 RA Shimoto M., Fukuhara Y., Itoh K., Oshima A., Sakuraba H., Suzuki Y.;
 RT "Protective protein gene mutations in galactosialidosis.";
 RL J. Clin. Invest. 91:2393-2398 (1993).
 CC -!- FUNCTION: PROTECTIVE PROTEIN APPEARS TO BE ESSENTIAL FOR BOTH THE
 CC ACTIVITY OF BETA-GALACTOSIDASE AND NEURAMINIDASE, IT ASSOCIATES
 CC WITH THESE ENZYMES AND EXERTS A PROTECTIVE FUNCTION NECESSARY FOR
 CC THEIR STABILITY AND ACTIVITY. THIS PROTEIN IS ALSO A
 CC CARBOXYPEPTIDASE AND CAN DEAMIDATE TACHYKININS.
 CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
 CC broad specificity.
 CC -!- SUBUNIT: HETERODIMER OF A 32 kDa AND 20 kDa CHAINS HELD TOGETHER
 CC BY DISULFIDE BONDS.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- DISEASE: Defects in PPG8 are the cause of galactosialidosis
 CC [MIM:256540], an autosomal recessive disease.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

EMBL; M22960; AAA36476.1; -.
EMBL; AL008726; CAA15501.1; -.
EMBL; BC000597; AAH00597.1; -.
PIR; A31589; A31589.
PDB; 1IVY; 2I-APR-97.
MEROPS; S10.002; -.
GIMM; HGNC:9251; PPOB.
MIM; 256540; -.
DR GO; GO:0005783; C:endoplasmic reticulum; TAS.
DR GO; GO:0004180; F:carboxypeptidase activity; TAS.
DR GO; GO:0008047; F:enzyme activator activity; TAS.
DR GO; GO:0005478; F:intracellular transporter activity; TAS.
DR GO; GO:0006886; F:intracellular protein transport; TAS.
DR InterPro; IPR000379; Ser esters site.
DR InterPro; IPR001563; Serine carbpept.
PFam; PF00450; serine carbpept; 1.
PRINTS; PR00724; CR50XPTASEC.
ProDom; PD01189; Serine carbpept; 2.
PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal; Lysosome;
KW 3D-structure; Disease mutation.
FT SIGNAL 1 28
FT CHAIN 29 480
FT CHAIN 29 480
FT CHAIN 327 480
FT ACT SITE 178 178
FT ACT SITE 400 400
FT ACT SITE 457 457
FT ACT SITE 457 457
FT DISULFID 88 362
FT DISULFID 240 256
FT DISULFID 241 246
FT DISULFID 281 331
FT CARBOHYD 145 145
FT CARBOHYD 333 333
FT VARIANT 49 49
FT VARIANT 65 65
FT VARIANT 90 90
FT VARIANT 249 249
FT VARIANT 395 395
FT VARIANT 440 440
FT MUTAGEN 178
FT MUTAGEN 457
FT CONFLICT 19 19
FT CONFLICT 56 56
FT HELIX 31 33
FT STRAND 34 34
FT TURN 39 40
FT STRAND 49 55
FT TURN 58 59
FT STRAND 60 67
FT TURN 73 75
FT STRAND 78 82
FT TURN 85 87
FT STRAND 88 88

```

Query Match 1.4%; Score 8; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 6;

```

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 100 LWLNGGPG 107
DB 80 LWLNGGPG 87
|||||
|||||

RESULT 10
CBP1_HORVU
ID CBP1_HORVU STANDARD; PRT; 499 AA.
AC P07519; P07520;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serine carboxypeptidase I precursor (EC 3.4.16.5) (Carboxypeptidase
DE C) (CP-MI).
GN CBP1 OR CXP;1.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RN [1] SEQUENCE FROM N.A.
RC TISSUE=Aleurone;
RA Rocher A., Lok F., Cameron-Mills V., von Wettstein D.;
RL Submitted (DSC-1996) to the EMBL/GenBank/DBJ databases.
[2]
RN [2] SEQUENCE OF 88-499 FROM N.A.
RP MEDLINE=88298749; PubMed=3403516;
RA Doan N.P., Fincher G.B.;
RT "The A- and B-chains of carboxypeptidase I from germinated barley
RL originate from a single precursor polypeptide.";
RN [3]
RN [3] J. Biol. Chem. 263:11106-11110(1988).
RP SEQUENCE OF 31-296 AND 352-499.
RA Seerensen S.B., Breddam K., Svendsen I.;
RT "Primary structure of carboxypeptidase I from malted barley.";
RL Carlsberg Res. Commun. 51:475-485(1986).
CC -!- FUNCTION: MAY BE INVOLVED IN THE DEGRADATION OF SMALL PEPTIDES (2-
CC 5 RESIDUES) OR IN THE DEGRADATION OF STORAGE PROTEINS IN THE
CC EMBRYO.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
CC broad specificity.
CC -!- SUBUNIT: CARBOXYPEPTIDASE I IS A DIMER, WHERE EACH MONOMER IS
CC COMPOSED OF TWO CHAINS LINKED BY DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: SECRETED INTO THE ENDOSPERM.
CC -!- DEVELOPMENTAL STAGE: AFTER ONE DAY OF GERMINATION, MAINLY FOUND IN
CC THE SCUTELLUM OF THE DEVELOPING GRAIN; BARELY DETECTABLE AFTER
CC FOUR DAYS, AND ABSENT FROM THE MATURE GRAIN. A LOWER LEVEL OF
CC EXPRESSION IS SEEN IN THE ALEURONE BOTH DURING DEVELOPMENT AND
CC GERMINATION.
CC -!- PTM: THREE DISULFIDE BONDS ARE PRESENT.
CC -!- PTM: THE LINKER PEPTIDE IS ENDOPROTEOLYTICALLY EXCISED DURING
CC ENZYME MATURATION.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

EMBL; Y09603; CAA70816.1; -.
EMBL; J03897; AAA32940.1; -.
PIR; T05367; CPEHS.
HSSP; P08819; 1WHT.
DR MEROPS; S10.004; -.
DR InterPro; IPR000379; Ser esters site.
DR InterPro; IPR001563; Serine carbpept.
PFam; PF00450; serine carbpept; 1.

```


DR ProDom; P0001189; Serine_carbpept; 1.
 DR PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
 KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal;
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 77 POTENTIAL.
 FT CHAIN 78 341 SERINE CARBOXYPEPTIDASE II-3, CHAIN A.
 FT CHAIN 342 352 LINKER PEPTIDE (BY SIMILARITY)
 FT CHAIN 353 516 SERINE CARBOXYPEPTIDASE II-3, CHAIN B.
 FT ACT SITE 236 236 BY SIMILARITY.
 FT ACT SITE 427 427 BY SIMILARITY.
 FT ACT SITE 484 484 BY SIMILARITY.
 FT DISULFID 143 391 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 300 315 BY SIMILARITY.
 FT DISULFID 339 359 BY SIMILARITY.
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 516 AA; 55913 MW; D41AALC56CF8D355 CRC64;
 Query Match 1.4%; Score 8; DB 1; Length 516;
 Best Local Similarity 100.0%; Pred. No. 6.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 100 LNLNGGPG 107
 Db 135 LNLNGGPG 142
 RESULT 13
 CBPY_PICPA STANDARD; PRT; 523 AA.
 AC P52710;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Carboxypeptidase Y precursor (EC 3.4.16.5) (Carboxypeptidase YSCY).
 GN PRC1.
 OS Pichia pastoris (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4922;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 108-142.
 RC STRAIN=GTSL15;
 RX MEDLINE=96381245; PubMed=8789258;
 RA Chi H., Ohtani W., Okazaki N., Furuhata N., Ohmura T.;
 RT "Cloning and characterization of the Pichia pastoris PRC1 gene
 encoding carboxypeptidase Y";
 RL Yeast 12:31-40(1996).
 CC -!- FUNCTION: INVOLVED IN DEGRADATION OF SMALL PEPTIDES.
 CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
 broad specificity.
 CC -!- SUBCELLULAR LOCATION: LYSOSOME-LIKE VACUOLES.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X87987; CAA61240.1; -.
 DR PIR; S61713; S61713.
 DR HSSP; P00729; 1YSC.
 DR MEROPS; S10.001; -.
 DR InterPro; IPR000379; Ser_estrs.site.
 DR InterPro; IPR001563; Serine_carbpept.
 DR Pfam; PF00450; serine_carbpept; 1.

DR PRINTS; PR00724; CRBOXYPTASRC.
 DR ProDom; P0001189; Serine_carbpept; 1.
 DR PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
 DR PROSITE; PS00560; CARBOXYPEPT SER HIS; 1.
 KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 107 POTENTIAL.
 FT CHAIN 108 523 CARBOXYPEPTIDASE Y.
 FT ACT SITE 249 249 BY SIMILARITY.
 FT ACT SITE 441 441 BY SIMILARITY.
 FT ACT SITE 498 498 BY SIMILARITY.
 FT DISULFID 162 401 BY SIMILARITY.
 FT DISULFID 296 310 BY SIMILARITY.
 FT DISULFID 320 343 BY SIMILARITY.
 FT DISULFID 327 336 BY SIMILARITY.
 FT DISULFID 365 371 BY SIMILARITY.
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 484 484 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 523 AA; 59447 MW; CA6093BDE2E9D178 CRC64;
 Query Match 1.4%; Score 8; DB 1; Length 523;
 Best Local Similarity 100.0%; Pred. No. 6.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 100 LNLNGGPG 107
 Db 154 LNLNGGPG 161
 RESULT 14
 CBPY_YEAST STANDARD; PRT; 532 AA.
 AC P00729;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Carboxypeptidase Y precursor (EC 3.4.16.5) (Carboxypeptidase YSCY).
 GN PRC1 OR YMR297W.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=87131100; PubMed=3028649;
 RA Valls L.A., Hunter C.P., Rothman J.H., Stevens T.H.;
 RT "Protein sorting in yeast: the localization determinant of yeast
 vacuolar carboxypeptidase Y resides in the propeptide";
 RL Cell 48:887-897(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX PubMed=9169872;
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
 RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajadream M.A.,
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
 XIII";
 RL Nature 387:90-93(1997).
 RN [3]
 RP SEQUENCE OF 112-532.
 RA Svendsen I., Martin B.M., Viswanatha T., Johansen J.T.;
 RT "Amino acid sequence of carboxypeptidase Y. II. Peptides from
 enzymatic cleavages";
 RL Carlsberg Res. Commun. 47:15-27(1982).
 RN [4]
 RP REVISIONS, AND ACTIVE SITE SER-257.
 RA Breddam K., Svendsen I.;
 RT "Identification of methionyl and cysteinyl residues in the substrate
 binding site of carboxypeptidase Y";

FT TURN 467 468
FT HELIX 469 474
FT STRAND 478 481
FT TURN 483 485
FT STRAND 488 492
FT STRAND 494 495
FT TURN 496 497
FT STRAND 498 503
FT STRAND 504 505
FT HELIX 510 513
FT HELIX 515 526
FT TURN 527 529
SQ SEQUENCE 532 AA; 53802 MW; 7227F3489CBDD952 CRC64;

Query Match 1.4%; Score 8; DB 1; Length 532;
Best Local Similarity 100.0%; Pred. No. 6.6; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
Db 159 LWLNGGPG 166

RESULT 15

CBPY_CANAL STANDARD; PRT; 542 AA.
AC P30574;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carboxypeptidase Y precursor (EC 3.4.16.5) (Carboxypeptidase YSCY).
GN CPY1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_taxid=5476;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93051356; PubMed=1427093;
RA Mukhtar M., Logan D.A., Kauffer N.F.;
RT "The carboxypeptidase Y-encoding gene from Candida albicans and its
transcription during yeast-to-hyphae conversion."
RL Gene 121:173-177(1992).
CC -!- FUNCTION: INVOLVED IN DEGRADATION OF SMALL PEPTIDES.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a
broad specificity.
CC -!- SUBCELLULAR LOCATION: LYSOSOME-LIKE VACUOLES.
CC -!- INDUCTION: TRANSIENTLY DOWN-REGULATED DURING THE EARLY EVENTS OF
YEAST TO HYPAE CONVERSION.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M95182; AAA34326.2; -.
DR HSSP; P00729; 1CPY.
DR MEROPS; S10.001; -.
DR InterPro; IPR000379; Ser esters site.
DR InterPro; IPR001563; Serine carboxypept.
DR Pfam; PF00450; serine carboxypept; 1.
DR PRINTS; PR00724; CRBOXYPTASEC.
DR PRODOM; PD001189; Serine carboxypept; 1.
DR PROSITE; PS00131; CARBOXYPEPT SER SER; 1.
DR PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.
KW Hydrolase; Carboxypeptidase; Glycoprotein; Zymogen; Signal.
FT SIGNAL 1 21
FT PROPEP 22 127 POTENTIAL.
FT CHAIN 128 542 CARBOXYPEPTIDASE Y.

FT ACT_SITE 269 269 BY SIMILARITY.
FT ACT_SITE 461 461 BY SIMILARITY.
FT ACT_SITE 518 518 BY SIMILARITY.
FT BINDING 464 464 SUBSTRATE (BY SIMILARITY).
FT BINDING 519 519 SUBSTRATE (BY SIMILARITY).
FT DISULFID 182 421 BY SIMILARITY.
FT DISULFID 316 330 BY SIMILARITY.
FT DISULFID 340 363 BY SIMILARITY.
FT DISULFID 347 356 BY SIMILARITY.
FT DISULFID 385 391 BY SIMILARITY.
FT CARBOHYD 213 213 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 542 AA; 61044 MW; 7FA6B9F82F9D44AF CRC64;

Query Match 1.4%; Score 8; DB 1; Length 542;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
Db 174 LWLNGGPG 181

Search completed: November 21, 2003, 18:04:24
Job time : 18 secs

112 Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 18:02:12 ; Search time 21 Seconds
(without alignments)

2541.601 Million cell updates/sec

Title: US-09-712-338-2

Perfect score: 555

Sequence: 1 MRGYEFLSVLPVLAASWALP.....HTQSSVPLPTATSMSSVGMA 555

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	1.8	423	2 S38953	carboxypeptidase D
2	10	1.8	460	1 S51516	serine-type carbox
3	9	1.6	405	2 S70150	tniQ protein - Xan
4	9	1.6	405	2 T08521	tniQ protein homol
5	9	1.6	468	2 D84503	probable serine ca
6	9	1.6	479	2 G96501	probable serine ca
7	9	1.6	495	2 S32179	tniQ protein homol
8	9	1.6	508	1 S46008	probable carboxype
9	9	1.6	2105	2 T18968	probable serine-type
10	8	1.4	36	2 S78239	ycf32 protein - Od
11	8	1.4	166	2 E71040	hypothetical prote
12	8	1.4	264	2 T49055	serine carboxypept
13	8	1.4	270	2 B75259	hypothetical prote
14	8	1.4	423	1 A29639	carboxypeptidase D
15	8	1.4	425	2 F85360	SERINE CARBOXYPEPT
16	8	1.4	425	2 E84631	probable serine ca
17	8	1.4	447	2 H84772	probable serine ca
18	8	1.4	452	2 H84772	probable serine ca
19	8	1.4	456	2 H86406	probable serine ca
20	8	1.4	456	2 A85139	hypothetical prote
21	8	1.4	458	2 F84746	probable serine ca
22	8	1.4	465	2 G86244	SERINE CARBOXYPEPT
23	8	1.4	465	2 B86358	SERINE CARBOXYPEPT
24	8	1.4	470	2 T16606	probable serine ca
25	8	1.4	470	2 B96637	hypothetical prote
26	8	1.4	474	2 A35732	protective protein
27	8	1.4	476	1 T05701	carboxypeptidase D
28	8	1.4	480	2 A31589	carboxypeptidase C
29	8	1.4	480	2 T50511	serine-type carbox

30 8 1.4 499 1 CPBHS carboxypeptidase C
31 8 1.4 501 2 T49081 serine-type carbox
32 8 1.4 502 2 T49188 serin carboxypepti
33 8 1.4 510 1 S43516 carboxypeptidase C
34 8 1.4 510 2 T39601 serine carboxypept
35 8 1.4 512 2 T33463 probable serine ca
36 8 1.4 516 2 S44191 carboxypeptidase D
37 8 1.4 523 1 S61713 carboxypeptidase C
38 8 1.4 532 1 CPBYI carboxypeptidase C
39 8 1.4 542 1 JC1380 carboxypeptidase C
40 8 1.4 552 2 JC7666 serine-type carbox
41 8 1.4 574 2 T16230 hypothetical prote
42 8 1.4 1002 2 T43236 carboxypeptidase C
43 8 1.4 2338 2 T25810 hypothetical prote
44 7 1.3 25 2 S78343 hypothetical prote
45 7 1.3 62 2 T21953 hypothetical prote

ALIGNMENTS

RESULT 1

S38953
carboxypeptidase D (EC 3.4.16.6) - Penicillium janthinellum
N:Alternate names: carboxypeptidase S1
C:Species: Penicillium janthinellum
C:Date: 18-Feb-1994 #sequence_revision 19-Apr-1996 #text_change 31-Mar-2000
C:Accession: S38953
R:Svendsen, I.; Hofmann, T.; Endrizzi, J.; Remington, S.J.; Breddam, K.
FEBS Lett. 333, 39-43, 1993
A:Title: The primary structure of carboxypeptidase S1 from Penicillium janthinellum.
A:Reference number: S38953; MUID:94039747; PMID:8224168
A:Accession: S38953
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-423 <SVE>
C:Superfamily: serine carboxypeptidase
C:Keywords: hydrolase; serine carboxypeptidase

Query Match 1.8%; Score 10; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 TSYGGHYGP 236

Db 141 TSYGGHYGP 150

RESULT 2

S51516
serine-type carboxypeptidase (EC 3.4.16.-) Z precursor - Absidia zychae

C:Species: Absidia zychae

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Jun-2000

C:Accession: S51516; S78013; S78014

R:Lee, B.R.; Takeuchi, M.; Kobayashi, Y.

Curr. Genet. 27, 159-165, 1995

A:Title: Molecular cloning and sequence analysis of the scpZ gene encoding the serine ca

A:Reference number: S51516; MUID:95308538; PMID:7788719

A:Accession: S51516

A:Molecule type: DNA

A:Residues: 1-460 <LEE>

A:Cross-references: EMBL:D16519; NID:G556466; PIDN:BAA03966.1; PID:g995456

A:Experimental source: strain NRIC 1199

A:Accession: S78013

A:Molecule type: protein

A:Residues: 52-62;90-99;367-381 <LEE>

A:Accession: S78014

A:Molecule type: mRNA

A:Residues: 18-460 <LES>

C:Genetics:

A:Gene: scpZ

A:Introns: 64/3; 95/1; 107/2; 138/3; 234/2; 246/1; 270/1; 343/3; 353/2; 384/3; 437/3

C:Superfamily: serine carboxypeptidase

C;Keywords: glycoprotein; hydrolase; serine carboxypeptidase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-51/Domain: propeptide #status predicted <PRE>
F:52-460/Product: serine-type carboxypeptidase #status experimental <MAT>
F:40,46,119/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:186/Active site: Ser #status predicted

Query Match 1.8%; Score 10; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 486 RYERAGHEVP 495
Db 430 RYERAGHEVP 439
|||||

RESULT 3
S70150
tniQ protein - Xanthomonas sp.
C;Species: Xanthomonas sp.
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 11-Jan-2000
C;Accession: S70150
R;Kholodii, G.Y.; Windlin, S.Z.; Bass, I.A.; Yurieva, O.V.; Minakhina, S.V.; Nikiforov, M.I. Microbiol. 17, 1189-1200, 1995
A;Title: Four genes, two ends, and a res region are involved in transposition of Tn5053:
A;Reference number: S70140; MUID:96130850; PMID:8594337
A;Accession: S70150
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-405 <RHO>
A;Cross-references: EMBL:L40585; NID:g710572; PIDN:AA98331.1; PID:g710584
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
C;Genetics:
A;Gene: tniQ
A;Start codon: GTG
A;Superfamily: Xanthomonas tniQ protein
C;Keywords: transposition

Query Match 1.6%; Score 9; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARNPETA 96
Db 341 EARNPETA 349
|||||

RESULT 4
T08521
tniQ protein homolog - Enterobacter aerogenes plasmid R751
C;Species: Enterobacter aerogenes
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: T08521
R;Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.
Plasmid 36, 95-111, 1996
A;Title: Conservation of the genetic switch between replication and transfer genes of In
A;Reference number: Z16434; MUID:97118926; PMID:8954881
A;Accession: T08521
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-405 <THO>
A;Cross-references: EMBL:U67194; NID:g1572520; PIDN:AA64465.1; PID:g1572570
C;Genetics:
A;Genome: plasmid R751
C;Superfamily: Xanthomonas tniQ protein
C;Keywords: transposition

Query Match 1.6%; Score 9; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 EARNPETA 96
Db 341 EARNPETA 349
|||||

Db 341 EARNPETA 349
|||||

RESULT 5
D84503
probable serine carboxypeptidase II [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: D84503
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, I.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84503
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-468 <STO>
A;Cross-references: GB:AE002093; NID:g4733981; PIDN:AA28662.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g12480
A;Map position: 2
C;Superfamily: serine carboxypeptidase

Query Match 1.6%; Score 9; DB 2; Length 468;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWNGGPG 107
Db 77 TLWNGGPG 85
|||||

RESULT 6
G96501
probable serine carboxypeptidases [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: G96501
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G96501
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-479 <STO>
A;Cross-references: GB:AE005173; NID:g7523661; PIDN:AAF63101.1; GSPDB:GN00141
C;Genetics:
A;Gene: F28H19.5
A;Map position: 1
C;Superfamily: serine carboxypeptidase

Query Match 1.6%; Score 9; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWNGGPG 107
Db 83 TLWNGGPG 91
|||||

RESULT 7
S32179
tniQ protein homolog - Klebsiella pneumoniae

C;Species: Klebsiella pneumoniae
 C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 11-Jan-2000
 C;Accession: S32179
 R;Radstroem, P.; Sundstrom, L.; Swedberg, G.; Flensburg, J.; Skold, O.
 submitted to the EMBL Data Library, March 1993
 A;Description: The interons of R751 and fn21 are transposons related to the retroelement
 A;Reference number: S32177
 A;Accession: S32179
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-495 <RAD>
 A;Cross-references: EMBL:X72585; NID:g288626; PIDN:CAA51177.1; PID:g288629
 C;Superfamily: Xanthomonas tniQ protein
 C;Keywords: transposition

Query Match 1.6%; Score 9; DB 2; Length 495;
 Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0

QY 88 EARNPETA 96
 |||||
 Db 431 EARNPETA 439

RESULT 8
 S46008
 probable carboxypeptidase C (EC 3.4.16.5) YBR139w - Yeast (Saccharomyces cerevisiae)
 N;Alternate names: hypothetical protein YBR1015; hypothetical protein YBR139w
 C;Species: Saccharomyces cerevisiae
 C;Date: 26-Aug-1994 #sequence_revision 05-Jan-1996 #text_change 19-Apr-2002
 C;Accession: S46008; S46581
 R;Becam, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulski, M.
 submitted to the Protein Sequence Database, August 1994
 A;Reference number: S45995
 A;Accession: S46008
 A;Molecule type: DNA
 A;Residues: 1-508 <BEC>
 A;Cross-references: EMBL:Z36008; NID:g536435; PIDN:CAA85097.1; PID:g536436; GSPDB:GN0000
 A;Experimental source: strain S288C
 R;Becam, A.M.; Cullin, C.; Grzybowska, E.; Lacroite, F.; Nasr, F.; Ozier-Kalogeropoulos,
 Yeast 10(Suppl.A), S1-S11, 1994
 A;Title: The sequence of 29.7kb from the right arm of chromosome II reveals 13 complete
 A;Reference number: S46569; MUID:94378717; EMID:8091856
 A;Accession: S46581
 A;Molecule type: DNA
 A;Residues: 1-508 <BE2>
 A;Cross-references: EMBL:X75891; NID:g496856; PIDN:CAA53497.1; PID:g496869
 A;Experimental source: strain S288C
 C;Genetics:
 A;Cross-references: SGD:S0000343
 A;Map position: 2R
 A;Note: MIPS:YBR139w
 C;Superfamily: serine carboxypeptidase
 C;Keywords: hydrolase; serine carboxypeptidase; transmembrane protein
 F;6-22/Domain; transmembrane status predicted <TM>
 F;219,415,474/Active site: Ser, Asp, His #status predicted

Query Match 1.6%; Score 9; DB 1; Length 508;
 Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0

QY 147 QPLGVGFSY 155
 |||||
 Db 171 QPLGVGFSY 179

RESULT 9
 T18968
 probable serine-type carboxypeptidase (EC 3.4.16.-) Y16B4A.2 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Mar-2000
 C;Accession: T18968; T23145; T26477
 R;Thomas, K.

submitted to the EMBL Data Library, March 1996
 A;Reference number: Z19053
 A;Accession: T18968
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-2105 <WIL>
 A;Cross-references: EMBL:Z70203; PIDN:CAA94110.1; GSPDB:GN000028; CESP:Y16B4A.2
 A;Experimental source: clone C05G5
 R;Lloyd, C.
 submitted to the EMBL Data Library, March 1997
 A;Reference number: Z19697
 A;Accession: T23145
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-2105 <WI2>
 A;Cross-references: EMBL:Z93339; PIDN:CAB07544.1; GSPDB:GN000028; CESP:Y16B4A.2
 A;Experimental source: clone H40D05
 R;Wallis, J.
 submitted to the EMBL Data Library, June 1998
 A;Reference number: Z20220
 A;Accession: T26477
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-2105 <WI3>
 A;Cross-references: EMBL:AL023825; PIDN:CAA19443.1; GSPDB:GN000028; CESP:Y16B4A.2
 A;Experimental source: clone Y16B4A
 C;Comment: This protein has 4 repeats of sequence homologous to serine carboxypeptidase.
 C;Genetics:
 A;Gene: CESP:Y16B4A.2
 A;Map position: X
 A;Introns: 43/3; 67/2; 179/1; 208/3; 239/2; 281/2; 296/3; 346/2; 444/3; 503/3; 588/3; 68
 1664/3; 1752/2; 1794/3; 1834/3; 1882/2; 1905/2; 1942/3; 1982/3
 C;Keywords: duplication; hydrolase; serine carboxypeptidase

Query Match 1.6%; Score 9; DB 2; Length 2105;
 Best Local Similarity 100.0%; Pred. No. 6; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0

QY 100 LWLNGPGS 108
 |||||
 Db 549 LWLNGPGS 557

RESULT 10
 S78239
 Ycf32 protein - Odontella sinensis chloroplast
 C;Species: chloroplast Odontella sinensis
 C;Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 20-Jun-2000
 C;Accession: S78239; S78293
 R;Kowallik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.
 Plant Mol. Biol. Rep. 13, 336-342, 1995
 A;Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella sinensis
 A;Reference number: S78238
 A;Accession: S78239
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-36 <KOW>
 A;Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91612.1; PID:g1185129
 A;Genetics: 5GEN
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
 A;Note: 5' gene
 A;Accession: S78293
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-36 <KOF>
 A;Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91666.1; PID:g1185183
 A;Genetics: 3GEN
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
 A;Note: 3' gene (c)
 C;Genetics: <GEN1>
 A;Gene: ycf32
 A;Genome: chloroplast
 A;Note: gene located on inverted repeat Ira

Mon Nov 24 13:41:07 2003

us-09-712-338-2-oli.rpr

C:Genetics: <GEN2>
A:Gene: ycf32,
A:Genome: Chloroplast
A:Note: gene located on inverted repeat IRb
C:Superfamily: hypothetical protein ycf32
C:Keywords: chloroplast

Query Match 1.4%; Score 8; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 LVAASWAL 19
Db 13 LVAASWAL 20
|||||

RESULT 11
E71040
hypothetical protein PH1613 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: E71040
R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: E71040
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-166 <RAW>
A:Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30725.1; PID:g3258042
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1613

Query Match 1.4%; Score 8; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 465 PLKVGVE 472
Db 43 PLKVGVE 50
|||||

RESULT 12
T49055
serine carboxypeptidase-like protein - Arabidopsis thaliana
N:Alternate names: protein TSP19.190
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49055
R:Benes, V.; Wurmbach, E.; Drzonek, H.; Ansoorge, W.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z25014
A:Accession: T49055
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <BEN>
A:Cross-references: EMBL:AL163972; GSPDB:GN00061; ATSP:TSP19.190
A:Experimental source: cultivar Columbia; BAC clone TSP19
C:Genetics:
A:Gene: ATSP:TSP19.190
A:Map position: 3
A:Introns: 131/1; 166/1

Query Match 1.4%; Score 8; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 LWLNGGPG 107
|||||

Db 126 LWLNGGPG 133

RESULT 13
B75259
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: B75259
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Praser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioreistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75259
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-270 <WHI>
A:Cross-references: GB:AE002085; GB:AE000513; NID:g6460383; PIDN:AAF12107.1; PID:g6460383
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2564
A:Map position: 1

Query Match 1.4%; Score 8; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 195 LAEEAAWE 202
Db 166 LAEEAAWE 173
|||||

RESULT 14
A29639
carboxypeptidase D (EC 3.4.16.6) - wheat
N:Alternate names: serine-type carboxypeptidase II
C:Species: Triticum aestivum (common wheat)
C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999
C:Accession: A29639; C29639; B29639
R:Breddam, K.; Sorensen, S.B.; Svendsen, I.
Carlsberg Res. Commun. 52, 297-311, 1987
A:Title: Primary structure and enzymatic properties of carboxypeptidase II from wheat
A:Reference number: A90771
A:Accession: A29639
A:Molecule type: protein
A:Residues: 1-263 <BRE>
A:Note: A chain
A:Accession: C29639
A:Molecule type: protein
A:Residues: 4-263 <BR2>
A:Note: A' chain
A:Accession: B29639
A:Molecule type: protein
A:Residues: 264-423 <BR3>
A:Note: B chain
C:Complex: homodimer of two mature products, each of which is cleaved to contain an A-
C:Superfamily: serine carboxypeptidase
C:Keywords: disulfide bond; glycoprotein; homodimer; hydrolase; serine carboxypeptidase
F:1-263,264-423/Product: carboxypeptidase D #status experimental <MAT>
F:1-263/Domain: A chain #status experimental <CHA1>
F:4-263/Domain: A chain #status experimental <CHA1>
F:264-423/Domain: B chain #status experimental <CHB>
F:116,127,259,291,297,421/Binding site: carboxylate (Asn) (covalent) #status experime
F:158,340,392/Active site: Ser, Asp, His #status predicted

Query Match 1.4%; Score 8; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 LWLNGGPG 107
|||||

Db 57 LWLNGGPG 64

RESULT 15

F85360

SERINE CARBOXYPEPTIDASE II-like protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001

C;Accession: F85360

R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nature 402, 769-777, 1999

A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: F85360

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-425 <STO>

A;Cross-references: GB:NC_001268; NID:97269982; PIDN:CAB79799.1; GSPDB:GN00140

C;Genetics:

A;Gene: At4g30810

A;Map position: 4

C;Superfamily: serine carboxypeptidase

Query Match

1.4%; Score 8; DB 2; Length 425;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107

|||||||

Db 85 LWLNGGPG 92

Search completed: November 21, 2003, 18:05:49

Job time : 22 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 18:03:57 ; Search time 36 Seconds
(without alignments)
2814.459 Million cell updates/sec

Title: US-09-712-338-2
Perfect score: 555
Sequence: 1 MRGYEFLSVLPVAASWALP.....HTQSSVPLPTATSSVCGMA 555

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 666188 seqs, 182559486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	1.4	209	12 US-10-294-039-4	Sequence 4, Appli
2	8	1.4	406	12 US-10-194-919-35	Sequence 35, Appl
3	8	1.4	411	9 US-09-901-252-16	Sequence 16, Appl
4	8	1.4	421	9 US-09-420-785A-4	Sequence 4, Appli
5	8	1.4	421	9 US-09-901-252-15	Sequence 15, Appl
6	8	1.4	480	10 US-09-736-457-336	Sequence 36, App
7	8	1.4	480	10 US-09-902-941-336	Sequence 36, App
8	8	1.4	480	10 US-09-849-921-336	Sequence 36, App
9	8	1.4	480	11 US-09-476-300-336	Sequence 36, App
10	8	1.4	480	12 US-10-113-872-336	Sequence 36, App
11	8	1.4	480	14 US-10-084-018-8	Sequence 8, Appli
12	8	1.4	480	15 US-10-017-754-336	Sequence 36, App
13	8	1.4	492	9 US-09-925-302-774	Sequence 774, App
14	8	1.4	492	11 US-09-796-753-42	Sequence 42, Appl
15	8	1.4	2991	12 US-09-970-944-22	Sequence 22, Appl

16	8	1.4	4349	12 US-10-174-677-76	Sequence 76, Appli
17	8	1.4	4349	12 US-09-970-944-4	Sequence 4, Appli
18	8	1.4	4349	12 US-09-970-944-18	Sequence 18, Appl
19	8	1.4	4349	12 US-09-970-944-19	Sequence 19, Appl
20	8	1.4	4349	15 US-10-160-758-15	Sequence 15, Appl
21	7	1.3	16	12 US-10-158-825-96	Sequence 96, Appl
22	7	1.3	16	15 US-10-158-847-96	Sequence 96, Appl
23	7	1.3	17	15 US-10-062-710-131	Sequence 131, App
24	7	1.3	34	9 US-09-776-490-24	Sequence 24, Appl
25	7	1.3	34	9 US-09-776-491-24	Sequence 24, Appl
26	7	1.3	70	9 US-09-867-550-112	Sequence 112, App
27	7	1.3	71	9 US-09-864-761-40430	Sequence 40430, A
28	7	1.3	72	10 US-09-954-512-1	Sequence 1, Appli
29	7	1.3	77	9 US-09-864-761-41457	Sequence 41457, A
30	7	1.3	85	10 US-09-751-100B-54	Sequence 54, Appl
31	7	1.3	85	10 US-09-751-100B-55	Sequence 55, Appl
32	7	1.3	85	10 US-09-751-100B-56	Sequence 56, Appl
33	7	1.3	85	10 US-09-751-100B-57	Sequence 57, Appl
34	7	1.3	85	10 US-09-751-100B-58	Sequence 58, App
35	7	1.3	85	14 US-10-092-243A-16	Sequence 16, Appl
36	7	1.3	91	12 US-09-933-767-584	Sequence 584, App
37	7	1.3	91	15 US-10-023-282-584	Sequence 8, Appli
38	7	1.3	163	9 US-09-735-487-8	Sequence 8, Appli
39	7	1.3	163	12 US-10-342-188-8	Sequence 8, Appli
40	7	1.3	163	15 US-10-156-761-12744	Sequence 12744, A
41	7	1.3	193	10 US-09-738-626-3874	Sequence 3874, Ap
42	7	1.3	225	10 US-09-908-805B-8	Sequence 8, Appli
43	7	1.3	229	10 US-09-738-626-5578	Sequence 5578, Ap
44	7	1.3	271	9 US-09-901-884-4	Sequence 4, Appli
45	7	1.3	317	12 US-10-333-631-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-10-294-039-4
; Sequence 4, Application US/10294039
; Publication No. US20030134814A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph M.
; TITLE OF INVENTION: METHOD OF USING 18080, A HUMAN SERINE
; TITLE OF INVENTION: CARBOXYPEPTIDE FAMILY MEMBER
; FILE REFERENCE: MP101-230PIRM
; CURRENT APPLICATION NUMBER: US/10/294,039
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/338,587
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(209)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-294-039-4

Query Match 1.4%; Score 8; DB 12; Length 209;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGPG 107
Db 33 LWLNGPG 40

RESULT 2
US-10-194-919-35
; Sequence 35, Application US/10194919
; Publication No. US20030145350A1

```

; GENERAL INFORMATION:
; APPLICANT: GVS Gesellschaft fuer Erwerb und Verwertung von Sc
; APPLICANT: Spener, Friedrich
; APPLICANT: Abbadi, Amine
; APPLICANT: Brummel, Monika
; TITLE OF INVENTION: METHOD FOR INCREASING THE CONTENT OF FATTY ACIDS IN PLANTS AND
; FILE REFERENCE: 1617.018US1
; CURRENT APPLICATION NUMBER: US/10/194,919
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: PCT/EP01/00289
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: DE 100 00 978.6
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-194-919-35

Query Match      1.4%; Score 8; DB 12; Length 406;
Best Local Similarity 100.0%; Pred. No. 28;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      193  TNLAAEA 200
Db      121  TNLAAEA 128

RESULT 3
US-09-901-252-16
; Sequence 16, Application US/09901252
; Patent No. US20020026658A1
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clint
; TITLE OF INVENTION: Genes Encoding Sinapoylglucose:Malate Sinapoyltransferase and Met
; FILE REFERENCE: N1422-004
; CURRENT APPLICATION NUMBER: US/09/901,252
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216593
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Triticales
US-09-901-252-16

Query Match      1.4%; Score 8; DB 9; Length 411;
Best Local Similarity 100.0%; Pred. No. 28;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100  LWLNGGPG 107
Db      57  LWLNGGPG 64

RESULT 4
US-09-420-785A-4
; Sequence 4, Application US/09420785A
; Patent No. US20010010923A1
; GENERAL INFORMATION:
; APPLICANT: MORTENSEN, UFFE
; APPLICANT: OLESEN, KJELD
; APPLICANT: STENNICKE, HENNING
; APPLICANT: SORENSEN, STEEN B.
; APPLICANT: BREDDAM, KLAUS
; TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE
; FILE REFERENCE: 089187/0109

```

```

; CURRENT APPLICATION NUMBER: US/09/420,785A
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-420-785A-4

Query Match      1.4%; Score 8; DB 9; Length 421;
Best Local Similarity 100.0%; Pred. No. 29;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100  LWLNGGPG 107
Db      48  LWLNGGPG 55

RESULT 5
US-09-901-252-15
; Sequence 15, Application US/09901252
; Patent No. US20020026658A1
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clint
; TITLE OF INVENTION: Genes Encoding Sinapoylglucose:Malate Sinapoyltransferase and Me
; FILE REFERENCE: N1422-004
; CURRENT APPLICATION NUMBER: US/09/901,252
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216593
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Saccaromyces
US-09-901-252-15

Query Match      1.4%; Score 8; DB 9; Length 421;
Best Local Similarity 100.0%; Pred. No. 29;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100  LWLNGGPG 107
Db      48  LWLNGGPG 55

RESULT 6
US-09-736-457-336
; Sequence 336, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, AiJun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 336
; LENGTH: 480

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-336

Query Match 1.4%; Score 8; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
| | | | |
Db 80 LWLNGGPG 87

RESULT 7

US-09-902-941-336
; Sequence 336, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-941-336

Query Match 1.4%; Score 8; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
| | | | |
Db 80 LWLNGGPG 87

RESULT 8

US-09-849-626-336
; Sequence 336, Application US/09849626
; Publication No. US2002019769A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-849-626-336

Query Match 1.4%; Score 8; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
| | | | |
Db 80 LWLNGGPG 87

RESULT 9

US-09-476-300-336
; Sequence 336, Application US/09476300
; Publication No. US20030125245A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C3
; CURRENT APPLICATION NUMBER: US/09/476,300
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 785
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-476-300-336

Query Match 1.4%; Score 8; DB 11; Length 480;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
| | | | |
Db 80 LWLNGGPG 87

RESULT 10

US-10-113-872-336
; Sequence 336, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-113-872-336

Query Match 1.4%; Score 8; DB 12; Length 480;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
| | | | |
Db 80 LWLNGGPG 87

; APPLICANT: Marnerakis, Margarita
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Vedwick, Thomas S.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: McNabb, Andria
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.478C18
 ; CURRENT APPLICATION NUMBER: US/10/017,754
 ; CURRENT FILING DATE: 2001-10-29
 ; NUMBER OF SEQ ID NOS: 2004
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 336
 ; LENGTH: 480
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-017-754-336

Query Match 1.4%; Score 8; DB 15; Length 480;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
 Db 80 LWLNGGPG 87

RESULT 13
 US-09-925-302-774
 ; Sequence 774, Application US/09925302
 ; Patent No. US20020044941A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA104
 ; CURRENT APPLICATION NUMBER: US/09/925,302
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05918
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 896
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 774
 ; LENGTH: 492
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-925-302-774

Query Match 1.4%; Score 8; DB 9; Length 492;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
 Db 92 LWLNGGPG 99

RESULT 14
 US-09-796-753-42
 ; Sequence 42, Application US/09796753
 ; Publication No. US20030027998A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCarthy, Sean A.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
 ; FILE REFERENCE: 7853-227-999
 ; CURRENT APPLICATION NUMBER: US/09/796,753
 ; CURRENT FILING DATE: 2001-03-01
 ; PRIOR APPLICATION NUMBER: 09/183,175
 ; PRIOR FILING DATE: 1998-10-30
 ; PRIOR APPLICATION NUMBER: 09/223,094
 ; PRIOR FILING DATE: 1998-12-30

RESULT 11
 US-10-084-018-8
 ; Sequence 8, Application US/10084018
 ; Publication No. US20020160499A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Hawkins, Phillip R.
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Lai, Preeti
 ; APPLICANT: Goli, Surya K.
 ; TITLE OF INVENTION: NOVEL HUMAN SERINE
 ; TITLE OF INVENTION: CARBOXYPEPTIDASE
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/084,018
 ; FILING DATE: 25-Feb-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/299,689A
 ; FILING DATE: <Unknown>
 ; APPLICATION NUMBER: 08/828,488
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PP-0241 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 480 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 190283
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
 US-10-084-018-8

Query Match 1.4%; Score 8; DB 14; Length 480;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
 Db 80 LWLNGGPG 87

RESULT 12
 US-10-017-754-336
 ; Sequence 336, Application US/10017754
 ; Publication No. US20030054363A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Watanabe, Yoshihiro
 ; APPLICANT: Johnson, Jeffrey C.
 ; APPLICANT: Retter, Marc W.

/ PRIOR APPLICATION NUMBER: 09/223,546
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/224,246
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/259,388
/ PRIOR FILING DATE: 1999-02-26
/ PRIOR APPLICATION NUMBER: 60/122,458
/ PRIOR FILING DATE: 1999-03-01
/ PRIOR APPLICATION NUMBER: 09/312,359
/ PRIOR FILING DATE: 1999-05-14
/ PRIOR APPLICATION NUMBER: 09/336,536
/ PRIOR FILING DATE: 1999-06-18
/ PRIOR APPLICATION NUMBER: 09/342,687
/ PRIOR FILING DATE: 1999-06-29
/ PRIOR APPLICATION NUMBER: 09/345,464
/ PRIOR FILING DATE: 1999-06-30
/ PRIOR APPLICATION NUMBER: 09/365,164
/ PRIOR FILING DATE: 1999-07-30
/ PRIOR APPLICATION NUMBER: 09/399,723
/ PRIOR FILING DATE: 1999-09-20
/ PRIOR APPLICATION NUMBER: 09/409,634
/ PRIOR FILING DATE: 1999-09-30
/ PRIOR APPLICATION NUMBER: 09/471,179
/ PRIOR FILING DATE: 1999-12-23
/ PRIOR APPLICATION NUMBER: 09/474,071
/ PRIOR FILING DATE: 1999-12-29
/ PRIOR APPLICATION NUMBER: 09/474,072
/ PRIOR FILING DATE: 1999-12-29
/ PRIOR APPLICATION NUMBER: 09/514,010
/ PRIOR FILING DATE: 2000-02-25
/ PRIOR APPLICATION NUMBER: 09/516,745
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: 09/572,002
/ PRIOR FILING DATE: 2000-05-14
/ PRIOR APPLICATION NUMBER: 09/597,993
/ PRIOR FILING DATE: 2000-06-19
/ PRIOR APPLICATION NUMBER: 09/599,596
/ PRIOR FILING DATE: 2000-06-22
/ PRIOR APPLICATION NUMBER: 09/630,334
/ PRIOR FILING DATE: 2000-07-31
/ PRIOR APPLICATION NUMBER: 09/606,565
/ PRIOR FILING DATE: 2000-06-29
/ PRIOR APPLICATION NUMBER: 09/606,317
/ PRIOR FILING DATE: 2000-06-29
/ PRIOR APPLICATION NUMBER: 09/665,666
/ PRIOR FILING DATE: 2000-09-20
/ PRIOR APPLICATION NUMBER: 09/677,751
/ PRIOR FILING DATE: 2000-09-30
/ NUMBER OF SEQ ID NOS: 162
/ SEQ ID NO 42
/ LENGTH: 492
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-796-753-42

Query Match 1.4%; Score 8; DB 11; Length 492;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
Db 93 LWLNGGPG 100

RESULT 15

US-09-970-944-22
/ Sequence 22, Application US/09970944
/ Publication No. US20030204052A1
/ GENERAL INFORMATION:
/ APPLICANT: Heriman, John L
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Shimkets, Richard A
/ TITLE OF INVENTION: No. US20030204052A1e1 Proteins and Nucleic Acids Encoding Same an

/ TITLE OF INVENTION: Antibodies Directed Against these Proteins
/ FILE REFERENCE: 21402-138
/ CURRENT APPLICATION NUMBER: US/09/970,944
/ CURRENT FILING DATE: 2002-05-02
/ PRIOR APPLICATION NUMBER: 60/237,862
/ PRIOR FILING DATE: 2000-10-04
/ NUMBER OF SEQ ID NOS: 62
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 22
/ LENGTH: 2991
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-970-944-22

Query Match 1.4%; Score 8; DB 12; Length 2991;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 INPVTGVV 170
Db 1278 INPVTGVV 1285

Search completed: November 21, 2003, 18:07:11
Job time : 37 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 18:03:02 ; Search time 22 Seconds
(without alignments)
1067.388 Million cell updates/sec

Title: US-09-712-338-2
Perfect score: 555
Sequence: 1 MRGYEFLSVLPLVAASWALP.....HTQSSVPLTATSMSSVGMA 555

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	99.8	554	3	US-08-943-714-2
2	35	6.3	35	3	US-08-943-714-5
3	19	3.4	20	3	US-08-943-714-3
4	16	2.9	29	3	US-08-943-714-6
5	10	1.8	423	3	US-08-943-714-9
6	8	1.4	81	4	US-09-439-554-2
7	8	1.4	283	4	US-09-107-532h-5418
8	8	1.4	291	4	US-09-439-554-20
9	8	1.4	421	2	US-08-807-263-4
10	8	1.4	480	2	US-08-828-488-8
11	8	1.4	480	4	US-09-299-689A-8
12	8	1.4	480	4	US-09-702-705-336
13	8	1.4	480	4	US-09-736-457-336
14	8	1.4	491	1	US-09-640-305-4
15	8	1.4	491	1	US-08-360-673-4
16	8	1.4	532	2	US-08-899-324-33
17	8	1.4	532	3	US-08-329-892B-33
18	8	1.4	557	1	US-08-309-341-2
19	8	1.4	557	1	US-08-309-341-4
20	8	1.4	557	1	US-08-608-267-2
21	8	1.4	557	1	US-08-608-267-4
22	8	1.4	557	1	US-08-608-452-2
23	8	1.4	557	1	US-08-608-452-4
24	8	1.4	557	1	US-08-608-224-2
25	8	1.4	557	1	US-08-608-224-4
26	8	1.4	557	1	US-08-967-149-2
27	8	1.4	557	2	US-08-967-149-4

```

28 7 1.3 9 3 US-08-159-339A-528 Sequence 528, App
29 7 1.3 10 3 US-08-159-339A-549 Sequence 549, App
30 7 1.3 19 3 US-08-943-714-4 Sequence 4, Appli
31 7 1.3 34 3 US-08-810-009-24 Sequence 24, Appl
32 7 1.3 55 2 US-08-637-759B-348 Sequence 348, App
33 7 1.3 55 3 US-08-871-355A-348 Sequence 348, App
34 7 1.3 55 4 US-09-201-945-348 Sequence 54, Appl
35 7 1.3 85 3 US-08-894-173-54 Sequence 54, Appl
36 7 1.3 85 3 US-08-894-173-55 Sequence 56, Appl
37 7 1.3 85 3 US-08-894-173-57 Sequence 57, Appl
38 7 1.3 85 3 US-08-894-173-58 Sequence 58, Appl
39 7 1.3 85 3 US-08-398-193-54 Sequence 54, Appl
40 7 1.3 85 3 US-08-398-193-55 Sequence 55, Appl
41 7 1.3 85 3 US-08-398-193-56 Sequence 56, Appl
42 7 1.3 85 3 US-08-398-193-57 Sequence 57, Appl
43 7 1.3 85 3 US-08-398-193-58 Sequence 58, Appl
44 7 1.3 91 4 US-09-205-258-584 Sequence 584, App
45 7 1.3

```

ALIGNMENTS

```

RESULT 1
US-08-943-714-2
; Sequence 2, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berk, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dambmann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6187578C No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,714
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4990.200-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 554 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-943-714-2

```

Query Match 99.8%; Score 554; DB 3; Length 554;

FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6187578e
US-08-943-714-5

Query Match 6.3%; Score 35; DB 3; Length 35;
Best Local Similarity 100.0%; Pred. No. 4.4e-29;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 383 DAIGNVNTQSNNDVYAFQQTGDFVWPNFIEDL 417
Db 1 DAIGNVNTQSNNDVYAFQQTGDFVWPNFIEDL 35

RESULT 3
US-08-943-714-3
Sequence 3, Application US/08943714
Patent No. 6187578
GENERAL INFORMATION:
APPLICANT: Blinkovsky, Alexander
APPLICANT: Berka, Randy
APPLICANT: Rey, Michael
APPLICANT: Golightly, Elizabeth
APPLICANT: Klotz, Alan
APPLICANT: Mathisen, Thomas Erik
APPLICANT: Dammann, Claus
TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
TITLE OF INVENTION: Encoding Same
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,714
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear


```

; MOLECULE TYPE: No. 6187578e
US-08-943-714-3
Query Match          3.4%; Score 19; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 230 YGGHYGPAFFNHFYEQNER 248
Db 2 YGGHYGPAFFNHFYEQNER 20

RESULT 4
US-08-943-714-6
; Sequence 6, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berkka, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dammann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; TITLE OF INVENTION: Encoding Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,714
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4990.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 12
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-943-714-6

Query Match          2.9%; Score 16; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 349 YAFAGRGVYDIRPYD 364
Db 11 YAFAGRGVYDIRPYD 26

RESULT 5
US-08-943-714-9
; Sequence 9, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berkka, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dammann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; TITLE OF INVENTION: Encoding Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,714
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4990.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6187578e
US-08-943-714-6

Query Match          1.8%; Score 10; DB 3; Length 423;
Best Local Similarity 100.0%; Pred. No. 0.077;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 227 TSYGGHYGP 236
Db 141 TSYGGHYGP 150

RESULT 6
US-09-439-554-2
; Sequence 2, Application US/09439554
; Patent No. 6479733
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Jan Antoni
; APPLICANT: Odell, Joan T.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Famodu, Omolayo O.
; TITLE OF INVENTION: STEROL METABOLISM ENZYMES
; FILE REFERENCE: BB1114 US NA
; CURRENT APPLICATION NUMBER: US/09/439,554
; CURRENT FILING DATE: 1999-11-12
; EARLIER APPLICATION NUMBER: 60/108,351
; EARLIER FILING DATE: 1998-No. 6479733member-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 81

```

TYPE: PRT
ORGANISM: Zea mays
US-09-439-554-2

Query Match 1.4%; Score 8; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 207 FLUGLPSL 214
Db 29 FLUGLPSL 36

RESULT 7

US-09-107-532A-5418
Sequence 5418, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5418:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:

NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...283
SEQUENCE DESCRIPTION: SEQ ID NO: 5418:
US-09-107-532A-5418

Query Match 1.4%; Score 8; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 379 DSVMDAIG 386
Db 187 DSVMDAIG 194

RESULT 8

US-09-439-554-20
Sequence 20, Application US/09439554
Patent No. 6479733
GENERAL INFORMATION:
APPLICANT: Rafalski, Jan Antoni
APPLICANT: Odell, Joan T.
APPLICANT: Sakai, Hajime
APPLICANT: Thorpe, Catherine J.
APPLICANT: Kinney, Anthony J.
APPLICANT: Pamodu, Omelayo O.
TITLE OF INVENTION: STEROL METABOLISM ENZYMES
FILE REFERENCE: BB1114 US NA
CURRENT APPLICATION NUMBER: US/09/439,554
CURRENT FILING DATE: 1999-11-12
EARLIER APPLICATION NUMBER: 60/108,351
EARLIER FILING DATE: 1998-No. 6479733ember-13
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 20
LENGTH: 291
TYPE: PRT
ORGANISM: Zea mays
US-09-439-554-20

Query Match 1.4%; Score 8; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 207 FLUGLPSL 214
Db 61 FLUGLPSL 68

RESULT 9

US-08-807-263-4
Sequence 4, Application US/08807263C
Patent No. 5985627
GENERAL INFORMATION:
APPLICANT: Mortensen, Uffe
APPLICANT: Olesen, Kjeld
APPLICANT: Stennicke, Henning
APPLICANT: Sorensen, Steen B.
APPLICANT: Bredam, Klaus
TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE
FILE REFERENCE: 8648.71us01-no4
CURRENT APPLICATION NUMBER: US/08/807,263C
CURRENT FILING DATE: 1997-02-28
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 421
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-08-807-263-4

Query Match 1.4%; Score 8; DB 2; Length 421;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 LWLNGGPG 107
Db 48 LWLNGGPG 55

RESULT 10

US-08-828-488-8
Sequence 8, Application US/08828488
Patent No. 5925521
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hawkins, Phillip R.
APPLICANT: Hillman, Jennifer L.

```

; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN SERINE
; TITLE OF INVENTION: CARBOXYPEPTIDASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,488
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0241 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 190283
; US-08-828-488-8

```

```

Query Match 1.4%; Score 8; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
Db 80 LWLNGGPG 87

```

```

RESULT 11
US-09-299-689A-8
; Sequence 8, Application US/09299689A
; Patent No. 6379913
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN SERINE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

```

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,689A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/828,488
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0241 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 190283
; US-09-299-689A-8

```

```

Query Match 1.4%; Score 8; DB 4; Length 480;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
Db 80 LWLNGGPG 87

```

```

RESULT 12
US-09-702-705-336
; Sequence 336, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 336
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-702-705-336

```

```

Query Match 1.4%; Score 8; DB 4; Length 480;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 LWLNGGPG 107
Db 80 LWLNGGPG 87

RESULT 13
US-09-736-457-336

```

Sequence 336, Application US/09736457
 Patent No. 6509448
 GENERAL INFORMATION:
 APPLICANT: Wang, Tongtong
 APPLICANT: Bangur, Chaitanya S.
 APPLICANT: Lodes, Michael A.
 APPLICANT: Fanger, Gary
 APPLICANT: Vedvick, Tom
 APPLICANT: Carter, Darick
 APPLICANT: Retter, Marc
 APPLICANT: Mannion, Jane
 APPLICANT: Fan, Liqun
 APPLICANT: Wang, Aijun
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 FILE REFERENCE: 210121.478C15
 CURRENT APPLICATION NUMBER: US/09/736,457
 CURRENT FILING DATE: 2000-12-13
 NUMBER OF SEQ ID NOS: 1864
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 336
 LENGTH: 480
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-736-457-336

Query Match 1.4%; Score 8; DB 4; Length 480;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 LWLNGGPG 107
 Db 80 LWLNGGPG 87

RESULT 14
 US-09-640-305-4
 Sequence 4, Application US/09640305
 Patent No. RE37447
 GENERAL INFORMATION:
 APPLICANT: Fleer, Reinhard
 Fournier, Alain
 Yeh, Patrice
 TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
 PREPARATION AND USE
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Rd. 3C43
 CITY: Collegeville
 STATE: PA
 COUNTRY: USA
 ZIP: 19002
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/640,305
 FILING DATE: 16-Aug-2000
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/360,673
 FILING DATE: 06-FEB-1995
 APPLICATION NUMBER: WO PCT/FR93/00623
 FILING DATE: 23-JUN-1993
 APPLICATION NUMBER: FR 92/07785
 FILING DATE: 25-JUN-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, Julie K.
 REGISTRATION NUMBER: 38 619
 REFERENCE/DOCKET NUMBER: ST92040-US
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (610)454-3839
 TELEFAX: (610)454-3808
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 491 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 US-09-640-305-4

Query Match 1.4%; Score 8; DB 1; Length 491;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 LWLNGGPG 107
 Db 113 LWLNGGPG 120

RESULT 15
 US-08-360-673-4
 Sequence 4, Application US/08360673
 Patent No. 5679544
 GENERAL INFORMATION:
 APPLICANT: Fleer, Reinhard
 APPLICANT: Fournier, Alain
 APPLICANT: Yeh, Patrice
 TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
 TITLE OF INVENTION: PREPARATION AND USE
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Rd. 3C43
 CITY: Collegeville
 STATE: PA
 COUNTRY: USA
 ZIP: 19002
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/360,673
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/FR93/00623
 FILING DATE: 23-JUN-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 92/07785
 FILING DATE: 25-JUN-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, Julie K.
 REGISTRATION NUMBER: 38,619
 REFERENCE/DOCKET NUMBER: ST92040-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610)454-3839
 TELEFAX: (610)454-3808
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 491 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-360-673-4

Query Match 1.4%; Score 8; DB 1; Length 491;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 LWLNGGPG 107

RESULT 16
 US-08-360-673-4
 Sequence 4, Application US/08360673
 Patent No. 5679544
 GENERAL INFORMATION:
 APPLICANT: Fleer, Reinhard
 APPLICANT: Fournier, Alain
 APPLICANT: Yeh, Patrice
 TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
 TITLE OF INVENTION: PREPARATION AND USE
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Rd. 3C43
 CITY: Collegeville
 STATE: PA
 COUNTRY: USA
 ZIP: 19002
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/360,673
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/FR93/00623
 FILING DATE: 23-JUN-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 92/07785
 FILING DATE: 25-JUN-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, Julie K.
 REGISTRATION NUMBER: 38,619
 REFERENCE/DOCKET NUMBER: ST92040-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610)454-3839
 TELEFAX: (610)454-3808
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 491 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-360-673-4

Query Match 1.4%; Score 8; DB 1; Length 491;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 LWLNGGPG 107

Mon Nov 24 13:41:06 2003

us-09-712-338-2.oli.ra1

Page 7

Db 113 LWLGGPG 120

Search completed: November 21, 2003, 18:06:23
Job time : 23 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2003, 18:00:11 ; Search time 46 Seconds
(without alignments)
1915.071 Million cell updates/sec

Title: US-09-712-338-2
Perfect score: 555
Sequence: 1 MRGYEFLSVLPLVAASWALP.....HTQSSVPLPTATSSVGVMA 555

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A Geneseq 19Jun03:*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	99.8	554	19	AAW56099
2	35	6.3	35	19	AAW57040
3	24	4.3	551	23	ABR38864
4	19	3.4	20	19	AAW56100
5	16	2.9	29	19	AAW57041
6	10	1.8	526	23	ABR38817
7	9	1.6	178	22	AU46801
8	9	1.6	454	21	AAW17034
9	9	1.6	461	21	AAW17033

10	9	1.6	461	21	AAW49759
11	9	1.6	461	21	AAW49762
12	9	1.6	473	21	AAW17032
13	9	1.6	473	21	AAW49758
14	9	1.6	473	21	AAW49761
15	9	1.6	508	23	ABR38859
16	9	1.6	554	23	ABR38859
17	9	1.6	623	23	ABR38865
18	8	1.4	81	24	ABU07821
19	8	1.4	93	22	AAO07036
20	8	1.4	116	23	ABP03224
21	8	1.4	216	15	AAW72106
22	8	1.4	240	21	AAW16761
23	8	1.4	246	21	AAW16760
24	8	1.4	291	24	ABU07830
25	8	1.4	303	22	ABG06385
26	8	1.4	438	18	AAW15092
27	8	1.4	444	23	ABW54514
28	8	1.4	450	21	AAW29802
29	8	1.4	452	18	AAW15091
30	8	1.4	452	21	AAW05734
31	8	1.4	452	21	AAW06851
32	8	1.4	456	21	AAW05733
33	8	1.4	458	21	AAW29801
34	8	1.4	470	21	AAW29800
35	8	1.4	479	21	AAW23905
36	8	1.4	479	21	AAW47178
37	8	1.4	480	22	AAW76860
38	8	1.4	480	23	AAW85515
39	8	1.4	480	24	ABU69487
40	8	1.4	480	24	ABU66389
41	8	1.4	486	21	AAW23904
42	8	1.4	486	21	AAW47177
43	8	1.4	491	15	AAW48059
44	8	1.4	492	21	AAW58436
45	8	1.4	492	21	AAW01416

ALIGNMENTS

RESULT 1

AAW56099
ID AAW56099 standard; Protein; 554 AA.

XX AC AAW56099;

XX 27-AUG-1998 (first entry)

DT DT
XX A. oryzae ATCC20386 carboxypeptidase I protein.

DE Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous; food industry.

XX OS Aspergillus oryzae.

XX PH Key Location/Qualifiers
FT Peptide 1..18
FT /label= signal

FT Protein 19..554
FT /label= carboxypeptidase I

XX MO9814599-A1.

XX 09-APR-1998.

XX 03-OCT-1997; 97WO-US17977.

XX 27-NOV-1996; 96US-0757534.

XX 04-OCT-1996; 96US-0726880.

XX (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO-NORDISK AS.

Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
S. cerevisiae BAX-
A. niger serine ca
A. niger serine ca
Corn C-4 methyl st
Human polypeptide
Human ORFX protein
Bovine growth horm
Arabidopsis thalia
Arabidopsis thalia
Corn C-4 methyl st
Novel human diagno
Human protective p
Lactococcus lactis
Arabidopsis thalia
Human precursor pr
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Human lung tumour
Clone #18991 of lu
Human lung cancer
Lung cancer therap
Arabidopsis thalia
Arabidopsis thalia
Sequence of protea
Lung cancer associ
Mouse TANGO 176.

XX Berka R, Blinkovsky A, Brown K, Dambmann C, Golightly E;
 PI Klotz A, Mathisen TE, Rey M;
 XX WPI; 1998-240098/21.
 DR N-PSDB; AAV28620.
 XX Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous
 PT substrates, useful for improving flavour of foods
 XX Claim 1; Fig 3; 82pp; English.
 XX This sequence represents carboxypeptidase I from Aspergillus oryzae.
 CC This polypeptide has an optimal activity in the range of pH 3.0-7.5 at
 CC 25 deg. C, optimal activity in the range of 55-60 deg. C at pH 4, and a
 CC residual activity of at least 65.5% after 30 minutes at pH 4.0 and
 CC 60 deg. C. It also has the capacity to hydrolyse X from N-CBZ-Ala-X where
 CC N-CBZ is N-carboxybenzoyl and X is any amino acid. The carboxypeptidases
 CC can be used for obtaining hydrolysates (which can be enriched in free
 CC glutamic acid or peptide bound glutamic acid residues) from proteinaceous
 CC substrates. The carboxypeptidases can be used in flavour-improving
 CC compositions in the food industry. The products can also be used for the
 CC production of polypeptides free of carboxypeptidase activity.
 XX SQ Sequence 554 AA;
 Query Match 99.8%; Score 554; DB 19; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 554; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRGYRFLSVLPVAAASWALPGSTPASVGRRLPKNTGKTLTANNVIRYKEPGAEGV 60
 Db 1 MRGYRFLSVLPVAAASWALPGSTPASVGRRLPKNTGKTLTANNVIRYKEPGAEGV 60
 QY 61 CETTRGVKSGVVDTSPESTHFFWFEARHNPETAPITLWNGPGSDSLGLFRELGP 120
 Db 61 CETTRGVKSGVVDTSPESTHFFWFEARHNPETAPITLWNGPGSDSLGLFRELGP 120
 QY 121 CHVNSFTDDYINPHSNWEVSNLLFLSQPLGVGFSYSDTVGNSINPTGVVENSFFAGVQG 180
 Db 121 CHVNSFTDDYINPHSNWEVSNLLFLSQPLGVGFSYSDTVGNSINPTGVVENSFFAGVQG 180
 QY 181 RYPTIDATLIDTNLAEEAAWEILQGLSLGSLDSRVQSKDFSLWTESYGGHYGPAFN 240
 Db 181 RYPTIDATLIDTNLAEEAAWEILQGLSLGSLDSRVQSKDFSLWTESYGGHYGPAFN 240
 QY 241 HFEQNERIANGSVNGVQLNFNSLGIINGIIDEAIQAPYYPEFAVNNYTGIKAVNETVYN 300
 Db 241 HFEQNERIANGSVNGVQLNFNSLGIINGIIDEAIQAPYYPEFAVNNYTGIKAVNETVYN 300
 QY 301 YMKFANQMPNGCQDLISTCKQTNRTALADYALCAEATNMCRDNVEGYPYAFAGRGVYDIR 360
 Db 301 YMKFANQMPNGCQDLISTCKQTNRTALADYALCAEATNMCRDNVEGYPYAFAGRGVYDIR 360
 QY 361 HPYDDPTPPSYNKFALKDSVMDAIGVNYTNSNDVYFAFOQTGDFVWPNFIEDLEEI 420
 Db 361 HPYDDPTPPSYNKFALKDSVMDAIGVNYTNSNDVYFAFOQTGDFVWPNFIEDLEEI 420
 QY 421 LALPVRVSLTYGDADYICNWFVGQAVSLAANYSOAQAQFRSAGYTPPLKVGVEYGETREYG 480
 Db 421 LALPVRVSLTYGDADYICNWFVGQAVSLAANYSOAQAQFRSAGYTPPLKVGVEYGETREYG 480
 QY 481 NFSFTRVYAGHEVYQPIASLQFNRTIFGMDIAEGOKKIWPSYKNGTATATHTQSS 540
 Db 481 NFSFTRVYAGHEVYQPIASLQFNRTIFGMDIAEGOKKIWPSYKNGTATATHTQSS 540
 QY 541 VPLPTATSMSSVGM 554
 Db 541 VPLPTATSMSSVGM 554
 RESULT 2
 AAW57040

ID AAW57040 standard; Protein; 35 AA.
 XX AAW57040;
 AC 27-AUG-1998 (first entry)
 DT 27-AUG-1998 (first entry)
 XX A. oryzae ATCC20386 carboxypeptidase I peptide fragment #3.
 DE Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous;
 XX food industry.
 KW Aspergillus oryzae.
 OS WO9814599-A1.
 XX 09-APR-1998.
 XX 03-OCT-1997; 97WO-US17977.
 XX 27-NOV-1996; 96US-0757534.
 PR 04-OCT-1996; 96US-0726880.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 PA (NOVO) NOVO-NORDISK AS.
 XX Berka R, Blinkovsky A, Brown K, Dambmann C, Golightly E;
 PI Klotz A, Mathisen TE, Rey M;
 XX WPI; 1998-240098/21.
 XX Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous
 PT substrates, useful for improving flavour of foods
 XX Example 2; Page 42; 82pp; English.
 XX This sequence represents a peptide fragment from carboxypeptidase I
 CC isolated from Aspergillus oryzae and generated from cyanogen bromide
 CC cleavage. this polypeptide has an optimal activity in the range of pH
 CC 3.0-7.5 at 25 deg. C, optimal activity in the range of 55-60 deg. C at
 CC pH 4, and a residual activity of at least 65.5% after 30 minutes at pH
 CC 4.0 and 60 deg. C. It also has the capacity to hydrolyse X from
 CC N-CBZ-Ala-X where N-CBZ is N-carboxybenzoyl and X is any amino acid. The
 CC carboxypeptidases can be used for obtaining hydrolysates (which can be
 CC enriched in free glutamic acid or peptide bound glutamic acid residues)
 CC from proteinaceous substrates. The carboxypeptidases can be used in
 CC flavour-improving compositions in the food industry. The products can
 CC also be used for the production of polypeptides free of carboxypeptidase
 CC activity.
 XX SQ Sequence 35 AA;
 Query Match 6.3%; Score 35; DB 19; Length 35;
 Best Local Similarity 100.0%; Pred. No. 2.1e-27;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 383 DAIGVNIYQSNNDVYFAFOQTGDFVWPNFIEDL 417
 Db 1 DAIGVNIYQSNNDVYFAFOQTGDFVWPNFIEDL 35
 RESULT 3
 ABR38864
 ID ABR38864 standard; Protein; 551 AA.
 XX ABR38864;
 XX 24-APR-2003 (first entry)
 DT A. niger serine carboxypeptidase polypeptide #3.
 DE Protease; fungal infection; aspergillosis; food; tanning; detergent;
 KW protein solubility; viscosity; taste; texture; nutritional value;
 KW EC3.4.16.6.

XX OS Aspergillus niger.
 XX PN WO200268623-A2.
 XX XX
 XX PD 06-SEP-2002.
 XX XX
 XX PF 22-FEB-2002; 2002WO-EP01984.
 XX XX
 XX PR 23-FEB-2001; 2001EP-0200657.
 XX PR 23-FEB-2001; 2001EP-0200658.
 XX PR 23-FEB-2001; 2001EP-0200660.
 XX PR 26-FEB-2001; 2001EP-0200706.
 XX PR 26-FEB-2001; 2001EP-0200707.
 XX PR 26-FEB-2001; 2001EP-0200708.
 XX PR 26-FEB-2001; 2001EP-0200719.
 XX PR 28-MAR-2001; 2001EP-0000075.
 XX PR 28-MAR-2001; 2001EP-0000078.
 XX PR 28-MAR-2001; 2001EP-0000080.
 XX PR 28-MAR-2001; 2001EP-0000087.
 XX PR 28-MAR-2001; 2001EP-0000156.
 XX PR 21-MAY-2001; 2001EP-0000159.
 XX PR 21-MAY-2001; 2001EP-0000160.
 XX PR 21-MAY-2001; 2001EP-0000162.
 XX PR 21-MAY-2001; 2001EP-0000165.
 XX PR 21-MAY-2001; 2001EP-0000166.
 XX PR 21-MAY-2001; 2001EP-0000168.
 XX PR 21-JUN-2001; 2001EP-0000240.
 XX PR 21-JUN-2001; 2001EP-0000242.
 XX PR 21-JUN-2001; 2001EP-0000244.
 XX PR 21-JUN-2001; 2001EP-0000246.
 XX PR 12-JUL-2001; 2001EP-0000280.
 XX PR 12-JUL-2001; 2001EP-0000285.
 XX PR 30-JUL-2001; 2001EP-0000323.
 XX PR 30-JUL-2001; 2001EP-0000327.
 XX PR 02-AUG-2001; 2001EP-0000341.
 XX PR 02-AUG-2001; 2001EP-0000342.
 XX PR 02-AUG-2001; 2001EP-0000343.
 XX PR 02-AUG-2001; 2001EP-0000344.
 XX PR 09-AUG-2001; 2001EP-0000357.
 XX PR 16-AUG-2001; 2001EP-0000374.
 XX PR 16-AUG-2001; 2001EP-0000377.
 XX PR 20-SEP-2001; 2001EP-0000478.
 XX PR 20-SEP-2001; 2001EP-0000483.
 XX PR 22-OCT-2001; 2001EP-0000552.
 XX PR 22-OCT-2001; 2001EP-0000553.
 XX PR 22-OCT-2001; 2001EP-0000554.
 XX PR 22-OCT-2001; 2001EP-0000556.
 XX PR 22-OCT-2001; 2001EP-0000557.
 XX PR 15-NOV-2001; 2001EP-0004464.
 XX PR 21-DEC-2001; 2001EP-0005117.
 XX XX
 XX PA (STAM) DSM NV.
 XX XX
 XX PI Edens L, Van Dijk AA, Krubasik P, Albermann K, Stock A, Kimpel E;
 XX PI Klugbauer S, Wagner C, Fritz A, Von Gustedt W, Heinrich O;
 XX PI Maier D, Spreafico F, Folkers U, Hopper S, Kemmer W, Tan P;
 XX PI Stiebler J, Albarg R;
 XX XX
 XX DR WPI; 2002-723203/78.
 XX DR P-PSDB; AB278231, AB278288.
 XX XX
 XX PT Novel isolated protease polypeptide useful in laboratory, clinical,
 XX PT pharmaceutical, chemical, diagnostic, personal care and industrial
 XX PT applications -
 XX XX
 XX PS Claim 13; Page 379-382; 394pp; English.
 XX XX
 XX CC The invention relates to a novel isolated protease polypeptide. A
 XX CC polypeptide or polynucleotide of the invention is useful for diagnosing a
 XX CC fungal infection such as aspergillosis, or as a query sequence to perform

a search against public databases. A polypeptide of the invention is
 useful in a selected number of industrial or pharmaceutical processes, in
 laboratory or clinical processes, in food industry (baking, brewing,
 cheese manufacturing, meat tenderising), in tanning industry and in the
 manufacture of biological detergents. A polypeptide may also be useful
 for improving protein solubility, extraction yields, viscosity or taste,
 texture, nutritional value, minimising of antigenicity or
 anti-nutritional factors, colour or functionality as well as processing
 aspects like filterability of the proteinaceous raw material. The
 sequences shown in ABR38313-ABR38869 represent the A. niger proteases of
 the invention.

XX SQ Sequence 551 AA;
 Query Match 4.3%; Score 24; DB 23; length 551;
 Best Local Similarity 100.0%; Pred. No. 3.6e-15;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWLNGGPGSDSLGLFEEYGPCH 122
 |||||
 Db 99 TLWLNGGPGSDSLGLFEEYGPCH 122

RESULT 4
 AAW56100
 ID AAW56100 standard; Protein; 20 AA.
 XX AC AAW56100;
 XX DT 27-AUG-1998 (first entry)
 XX DE A. oryzae ATCC20386 carboxypeptidase I peptide fragment #1.
 XX KW Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous;
 XX OS food industry.
 XX OS Aspergillus oryzae.
 XX PH Key Location/Qualifiers
 FT Modified-site 1
 FT /label= unknown
 XX WO9814599-A1.
 XX PD 09-APR-1998.
 XX PF 03-OCT-1997; 97WO-US17977.
 XX PR 27-NOV-1996; 96US-0757534.
 XX PR 04-OCT-1996; 96US-0726880.
 XX PA (NOVO) NOVO NORDISK BIOTECH INC.
 XX PA (NOVO) NOVO-NORDISK AS.
 XX XX
 XX PI Berka R, Blinkovsky A, Brown K, Danbmann C, Golightly E;
 XX PI Klotz A, Mathisen TE, Rey M;
 XX DR WPI; 1998-240098/21.
 XX XX
 XX PT Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous
 XX PT substrates, useful for improving flavour of foods
 XX PS Example 2; Page 41; 82pp; English.
 XX XX
 XX CC This sequence represents an N-terminal peptide from carboxypeptidase I
 CC isolated from Aspergillus oryzae. This polypeptide has an optimal
 CC activity in the range of pH 3.0-7.5 at 25 deg. C, optimal activity in
 CC the range of 55-60 deg. C at pH 4, and a residual activity of at least
 CC 65.5% after 30 minutes at pH 4.0 and 60 deg. C. It also has the capacity
 CC to hydrolyse X from N-CBZ-Ala-X where N-CBZ is N-carbobenzoxy and X is
 CC any amino acid. The carboxypeptidases can be used for obtaining
 CC hydrolysates (which can be enriched in free glutamic acid or peptide
 CC bound glutamic acid residues) from proteinaceous substrates. The

CC carboxypeptidases can be used in flavour-improving compositions in the
 CC food industry. The products can also be used for the production of
 CC polypeptides free of carboxypeptidase activity.

XX SQ Sequence 20 AA;

Query Match 3.4%; Score 19; DB 19; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2e-11;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 YGCHYGPAFFNHFYQNER 248
 DB 2 YGCHYGPAFFNHFYQNER 20

RESULT 5

AAW57041
 ID AAW57041 standard; Protein; 29 AA.

XX AC AAW57041;

XX DT 27-AUG-1998 (first entry)

XX DE A. oryzae ATCC20386 carboxypeptidase I peptide fragment #4.

XX KW Carboxypeptidase I; flavour improving agent; hydrolysate; proteinaceous;
 food industry.

XX OS Aspergillus oryzae.

XX FH Key Location/Qualifiers

FT Modified-site 1 /label= unknown

FT Modified-site 10 /label= unknown

XX WO9814599-A1.

XX PD 09-APR-1998.

XX PF 03-OCT-1997; 97WO-US17977.

XX PR 27-NOV-1996; 96US-0757534.

XX PR 04-OCT-1996; 96US-0726880.

XX PA (NOVO) NOVO NORDISK BIOTECH INC.

XX PA (NOVO) NOVO-NORDISK AS.

XX PI Berka R, Blinkovsky A, Brown K, Dambmann C, Golightly E;

XX PI Klotz A, Mathisen TE, Rey M;

XX DR WPI; 1998-240098/21.

XX PT Carboxypeptidase from Aspergillus oryzae - hydrolyses proteinaceous
 PT substrates, useful for improving flavour of foods

XX PS Example 2; Page 42; 82pp; English.

XX CC This sequence represents a peptide fragment from carboxypeptidase I
 CC isolated from Aspergillus oryzae and generated from cyanogen bromide
 CC cleavage. This polypeptide has an optimal activity in the range of pH
 CC 3.0-7.5 at 25 deg. C, optimal activity in the range of 55-60 deg. C at
 CC pH 4, and a residual activity of at least 65.5% after 30 minutes at pH
 CC 4.0 and 60 deg. C. It also has the capacity to hydrolyse X from
 CC N-CBZ-Ala-X where N-CBZ is N-carboxybenzyl and X is any amino acid. The
 CC carboxypeptidases can be used for obtaining hydrolysates (which can be
 CC enriched in free glutamic acid or peptide bound glutamic acid residues)
 CC from proteinaceous substrates. The carboxypeptidases can be used in
 CC flavour-improving compositions in the food industry. The products can
 CC also be used for the production of polypeptides free of carboxypeptidase
 CC activity.

XX SQ Sequence 29 AA;

Query Match 2.9%; Score 16; DB 19; Length 29;
 Best Local Similarity 100.0%; Pred. No. 3.1e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 YAFAGRGVYDIRHPYD 364

DB 11 YAFAGRGVYDIRHPYD 26

RESULT 6

ABR38817

ID ABR38817 standard; Protein; 526 AA.

XX AC ABR38817;

XX DT 24-APR-2003 (first entry)

XX DE A. niger serine carboxypeptidase polypeptide #1.

XX KW Protease; fungal infection; aspergillosis; food; tanning; detergent;
 KW protein solubility; viscosity; taste; texture; nutritional value;
 KW EC3.4.16.6.

XX OS Aspergillus niger.

XX PN WO200268623-A2.

XX PD 06-SEP-2002.

XX PF 22-FEB-2002; 2002WO-EP01984.

XX PR 23-FEB-2001; 2001EP-0200657.

XX PR 23-FEB-2001; 2001EP-0200658.

XX PR 23-FEB-2001; 2001EP-0200660.

XX PR 26-FEB-2001; 2001EP-0200706.

XX PR 26-FEB-2001; 2001EP-0200707.

XX PR 26-FEB-2001; 2001EP-0200708.

XX PR 26-FEB-2001; 2001EP-0200719.

XX PR 28-MAR-2001; 2001EP-0000075.

XX PR 28-MAR-2001; 2001EP-0000078.

XX PR 28-MAR-2001; 2001EP-0000080.

XX PR 28-MAR-2001; 2001EP-0000087.

XX PR 28-MAR-2001; 2001EP-0000088.

XX PR 21-MAY-2001; 2001EP-0000156.

XX PR 21-MAY-2001; 2001EP-0000159.

XX PR 21-MAY-2001; 2001EP-0000160.

XX PR 21-MAY-2001; 2001EP-0000162.

XX PR 21-MAY-2001; 2001EP-0000165.

XX PR 21-MAY-2001; 2001EP-0000166.

XX PR 21-MAY-2001; 2001EP-0000168.

XX PR 21-JUN-2001; 2001EP-0000240.

XX PR 21-JUN-2001; 2001EP-0000242.

XX PR 21-JUN-2001; 2001EP-0000244.

XX PR 21-JUN-2001; 2001EP-0000246.

XX PR 12-JUL-2001; 2001EP-0000280.

XX PR 12-JUL-2001; 2001EP-0000285.

XX PR 30-JUL-2001; 2001EP-0000323.

XX PR 30-JUL-2001; 2001EP-0000327.

XX PR 02-AUG-2001; 2001EP-0000341.

XX PR 02-AUG-2001; 2001EP-0000342.

XX PR 02-AUG-2001; 2001EP-0000343.

XX PR 02-AUG-2001; 2001EP-0000344.

XX PR 09-AUG-2001; 2001EP-0000357.

XX PR 16-AUG-2001; 2001EP-0000377.

XX PR 16-AUG-2001; 2001EP-0000378.

XX PR 20-SEP-2001; 2001EP-0000483.

XX PR 20-SEP-2001; 2001EP-0000483.

XX PR 22-OCT-2001; 2001EP-0000552.

XX PR 22-OCT-2001; 2001EP-0000553.

XX PR 22-OCT-2001; 2001EP-0000554.

XX PR 22-OCT-2001; 2001EP-0000556.

XX PR 22-OCT-2001; 2001EP-0000557.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149910.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155119.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 22-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160988.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 9; DB 21; Length 454;
 Best Local Similarity 100.0%; Pred. No. 4.8;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 TLWLNGGPG 107

Db 59 TLWLNGGPG 67

RESULT 9

AAG17033
 ID AAG17033 standard; Protein; 461 AA.

XX AC AAG17033;

XX DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 17901.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 20-MAY-1999; 99US-0134941.
 PR 21-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137582.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 18-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149388.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155133.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155655.
PR 28-SEP-1999; 99US-0156459.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 9; DB 21; Length 461;

Best Local Similarity 100.0%; Pred. No. 4.8; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0;

QY 99 TLWLNGGPG 107
|||||||
Db 66 TLWLNGGPG 74

RESULT 10

AAG49759

ID AAG49759 standard; Protein; 461 AA.

XX AC AAG49759;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 62381.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 23-APR-1999; 99US-0130891.

XX PR 28-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 30-APR-1999; 99US-0132407.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140333.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

[illegible]

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149925.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159233.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match

Best Local Similarity 1.6%; Score 9; DB 21; Length 461;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWLNGGPG 107
|||
DB 66 TLWLNGGEG 74

RESULT 12

AAAG17032
ID AAAG17032 standard; Protein; 473 AA.

XX
AC AAAG17032;

XX
DT 17-OCT-2000 (first entry)

XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 17900.

XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX
PN EP1033405-A2.

XX
PD 06-SEP-2000.

XX
EF 25-FEB-2000; 2000EP-0301439.

XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.

us-09-712-338-2.oli.rag

Mon Nov 24 13:41:05 2003

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142254.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145911.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149226.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

```
PR 29-OCT-1999; 99US-01G2142.
Query Match 1.6%; Score 9; DB 21; Length 473;
Best Local Similarity 100.0%; Pred.No. 4.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 99 TLWLNGGPG 107
Db 78 TLWLNGGPG 86

RESULT 13
AAG49758
ID AAG49758 standard; Protein; 473 AA.
XX
AC AAG49758;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62980.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123588.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
```

DB	78 TLWNGGPG 86																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
----	----------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 9; DB 21; Length 473;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 TLWLNGPG 107
|||
Db 78 TLWLNGPG 86
|||

RESULT 15
ABG93049
ID ABG93049 standard; Protein; 508 AA.

XX ABG93049;
XX
XX 21-NOV-2002 (first entry)
XX
XX DE
XX DE
XX S. cerevisiae BAX-associated protein fragment SEQ ID 56.
XX Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
KW neurodegeneration; cell death.
XX
XX Saccharomyces cerevisiae.
XX
XX WO200264766-A2.
XX
XX 22-AUG-2002.
XX
XX 21-DEC-2001; 2001WO-BF15398.
XX
XX 22-DEC-2000; 2000EP-0870318.
PR 04-JAN-2001; 2001EP-0870002.
PR 09-JAN-2001; 2001EP-0870003.
XX
XX (JANC) JANSSEN PHARM NV.
XX PA
XX
XX Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX
XX WPI; 2002-667002/71.
DR N-PSDB; ABQ76315.
XX
XX New isolated nucleic acid representing a synthetic Bax gene, useful as
PT medicament for treating, preventing and/or alleviating yeast or fungal
PT infections or proliferative disorders, or for preventing apoptosis in
PT certain diseases
XX
XX Claim 36; Figure 1; 344pp; English.
XX
XX This invention describes a novel nucleic acid representing a synthetic
CC Bax gene. The Bax gene of the invention is useful for identifying
CC Bax-resistant yeast or fungi, identifying, or obtaining and identifying
CC Candida spp. sequences that are differentially expressed in a pathway
CC eventually leading to programmed cell death or identifying inhibitors or
CC inhibitor sequences of Bax-induced cell death. The products of the
CC invention have cytostatic, fungicide; immunosuppressive, virucide and
CC vasotropic activity and can be used in vaccines or for gene therapy. The
CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
CC antisense molecules and antibodies are useful as medicaments or in
CC preparing a medicament for treating, preventing and/or alleviating
CC diseases associated with yeast or fungi or proliferative disorders, such
CC as cancer, or for preventing apoptosis in certain diseases. The compounds
CC or polypeptides, or the genetically modified organism are useful for
CC preparing a medicament for modifying the endogenous flora of humans and
CC other mammals. The vaccine is useful for immunising against yeast or
CC fungal infections. Apoptosis-related diseases include autoimmune disease,
CC ischaemia, diseases related with viral infections or neurodegenerations.
CC This sequence represents a polypeptide associated with the Bax gene
CC described in the disclosure of the invention.
XX
XX Sequence 508 AA;
SQ
Query Match 1.6%; Score 9; DB 23; Length 508;
Best Local Similarity 100.0%; Pred.No. 5.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 147 QPLGVGFSY 155
Db 171 QPLGVGFSY 179
Search completed: November 21, 2003, 18:03:54
Job time : 47 secs